

Help Logout Interrupt

Main Menu Search Form Posting Counts Show 8 Numbers Edit 8 Numbers Preferences Cases

Search Results -

| | |
|-----------|-----------|
| Terms | Documents |
| 19 and 15 | 0 |

Database:

- US Patents Full-Text Database
- US Pre-Grant Publication Full-Text Database
- JPO Abstracts Database
- EPO Abstracts Database
- Derwent World Patents Index
- IBM Technical Disclosure Bulletins

Search:

L13

Refine Search

Recall Text

Clear

Search History

DATE: Thursday, April 24, 2003 Printable Copy Create Case

Hit Count Set Name

result set

Set Name Query

side by side

DB=USPT,PGPB,PLUR=YES,OP=AND

| | | | |
|-----|--|-------|-----|
| L13 | 19 and 15 | 0 | L13 |
| L12 | 19 and 14 | 0 | L12 |
| L11 | 19 and 13 | 0 | L11 |
| L10 | L9 and 11 | 0 | L10 |
| L9 | monophosphoryl adj lipid adj a or aluminum adj salt or qs21 or isa adj 720 or saf or iscoms or mf-59 or sbas-2 or sbas-4 or detox or rc-529 or glucosaminide-4-phosphate | 15240 | L9 |
| L8 | adjuvant and 13 | 0 | L8 |
| L7 | adjuvant and 14 | 0 | L7 |
| L6 | adjuvant and L5 | 0 | L6 |
| L5 | 5773246.pn. | 1 | L5 |
| L4 | 5444149.pn. | 1 | L4 |
| L3 | 5525495.pn. | 1 | L3 |
| L2 | adjuvant and L1 | 0 | L2 |
| L1 | 6255055.pn. | 1 | L1 |

END OF SEARCH HISTORY

=> d his

(FILE 'HOME' ENTERED AT 15:48:12 ON 24 APR 2003)

FILE 'MEDLINE, CAPLUS, BIOSIS, SCISEARCH' ENTERED AT 15:48:26 ON 24 APR

2003

1253 S KOC

0 S MUELLER-PALTASCH AND L1

71 S L1 AND CANCER
36 DUP REM L3 (35 DUPLICATES REMOVED)

=> d au ti so 1-36 14

L4 ANSWER 1 OF 36 MEDLINE
AU Wang T; Fan L; Watanabe Y; McNeill P D; Moulton G G; Bangur C; Fanger G R;
TI Okada M; Inoue Y; Persing D H; Reed S G
L523S, an RNA-binding protein as a potential therapeutic target for lung
cancer.
SO BRITISH JOURNAL OF CANCER, (2003 Mar 24) 88 (6) 887-94.
Journal code: 0370635. ISSN: 0007-0920.

L4 ANSWER 2 OF 36 MEDLINE
AU Mueller F; Bommer M; Lacher U; Ruhland C; Stage V; Adler G; Gress T M;
TI Seufferlein T
KOC is a novel molecular indicator of malignancy.
SO BRITISH JOURNAL OF CANCER, (2003 Mar 10) 88 (5) 699-701.
Journal code: 0370635. ISSN: 0007-0920.

L4 ANSWER 3 OF 36 MEDLINE
AU Zhang Jian-Ying; Castano Carlos A; Peng Xuan-Xian; Kozlowski James A; Chan
TI Enhancement of Antibody Detection in Cancer Using Panel of
Recombinant Tumor-associated Antigens.

SO CANCER EPIDEMIOLOGY, BIOMARKERS AND PREVENTION, (2003 Feb) 12 (2) 136-43.
Journal code: 9200608. ISSN: 1055-9965.

L4 ANSWER 4 OF 36 CAPLUS COPYRIGHT 2003 ACS
IN Mack, David H.; Gish, Kurt C.
TI Gene expression profiles useful for diagnosis of human ovarian cancer and
screening for modulators of ovarian cancer
SO PCT Int. Appl., 332 pp.
CODEN: PIXXD2

L4 ANSWER 5 OF 36 CAPLUS COPYRIGHT 2003 ACS
IN Aziz, Natasha; Murray, Richard
TI Gene expression profile in human lung cancer and its use in diagnosis and
screening for modulators of lung cancer
SO PCT Int. Appl., 453 pp.
CODEN: PIXXD2

L4 ANSWER 6 OF 36 CAPLUS COPYRIGHT 2003 ACS
IN Murray, Richard; Glynn, Richard; Watson, Susan R.; Aziz, Natasha
TI Gene expression profiles associated with angiogenesis and their use in
diagnosis and screening for angiogenesis modulators
SO PCT Int. Appl., 291 pp.
CODEN: PIXXD2

L4 ANSWER 7 OF 36 CAPLUS COPYRIGHT 2003 ACS
IN Mueller, Friederike; Gress, Thomas; Adler, Guido
TI Medicament comprising a DNA sequence, which codes for the RNA-binding
KOC protein, and comprising a KOC protein or a DNA
sequence of the KOC promoter
SO PCT Int. Appl., 75 pp.

CODEN: PIXXD2

L4 ANSWER 8 OF 36 MEDLINE
AU Hoo Linda Soo; Zhang Jinying Y; Chan Edward K L
TI Cloning and characterization of a novel 90 kDa 'companion' auto-antigen of p62 overexpressed in cancer.
SO ONCOGENE, (2002 Jul 25) 21 (32) 5006-15.
Journal code: 8711562. ISSN: 0950-9232.

L4 ANSWER 9 OF 36 SCISEARCH COPYRIGHT 2003 ISI (R)
AU Itamochi H (Reprint); Kigawa J; Sultana H; Iba T; Akeshima R; Kamazawa S; Kanamori Y; Terakawa N
TI Sensitivity to anticancer agents and resistance mechanisms in clear cell carcinoma of the ovary
SO JAPANESE JOURNAL OF CANCER RESEARCH, (JUN 2002) Vol. 93, No. 6, pp. 723-728.
Publisher: BUSINESS CENTER ACADEMIC SOCIETIES JAPAN, 5-16-9 HONKOMAGOME, BUNKYO-KU, TOKYO, 113-8633, JAPAN.
ISSN: 0910-5050.

L4 ANSWER 10 OF 36 CAPLUS COPYRIGHT 2003 ACS
AU Wang, Shanying; Liu, Long; Wang, Lin; Lian, Qizhou; Li, Qifu; Peng, Xuanxian
TI Auto-antibody to Koc in patients with cancers
SO Zhongguo Mianyixue Zazhi (2002), 18(1), 30-32
CODEN: ZMZAEE; ISSN: 1000-484X

L4 ANSWER 11 OF 36 MEDLINE
AU Kiyozuka Y; Nakagawa H; Senzaki H; Uemura Y; Adachi S; Teramoto Y; Matsuyama T; Bessho K; Tsubura A
TI Bone morphogenetic protein-2 and type IV collagen expression in psammoma body forming ovarian cancer.
SO ANTICANCER RESEARCH, (2001 May-Jun) 21 (3B) 1723-30.
Journal code: 8102988. ISSN: 0250-7005.

L4 ANSWER 12 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AU Tan, Eng M. (1)
TI Autoantibodies as reporters identifying aberrant cellular mechanisms in tumorigenesis.
SO Journal of Clinical Investigation, (November, 2001) Vol. 108, No. 10, pp. 1411-1415. print.
ISSN: 0021-9738.

L4 ANSWER 13 OF 36 CAPLUS COPYRIGHT 2003 ACS
AU Martinez, Eduardo J; Corey, E. J.; Owa, Takashi
TI Antitumor activity- and gene expression-based profiling of ecteinascidin Et 743 and phtalascidin Pt 650
SO Chemistry & Biology (2001), 8(12), 1151-1160
CODEN: CBOLE2; ISSN: 1074-5521

L4 ANSWER 14 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AU Kohno, T. (1); Suzuki, M. (1); Mizukami, H.; Saga, Y. (1); Takei, Y. (1); Shimpō, M.; Ohwada, M. (1); Kume, A.; Ozawa, K. (1); Sato, I. (1)
TI Expression of interleukin-10 inhibits angiogenesis and tumor growth in ovarian cancer.
SO European Journal of Cancer, (October, 2001) Vol. 37, No. Supplement 6, pp. S325. http://www.elsevier.com/locate/ejca. print.
Meeting Info.: 11th European Cancer Conference Lisbon, Portugal October 21-25, 2001
ISSN: 0959-8049.

L4 ANSWER 15 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AU Bommer, Martin (1); Seufferlein, Thomas; Gress, Thomas; Adler, Guido; Doeberner, Hartmut (1); Mueller, Friederike

TI Detection of CNS involvement in patients with acute leukemia and Non-Hodgkin's lymphoma using RT-PCR for KOC-expression. Blood, (November 16, 2001) Vol. 98, No. 11 Part 2, pp. 157b. <http://www.bloodjournal.org/>. print. Meeting Info.: 43rd Annual Meeting of the American Society of Hematology, Part 2 Orlando, Florida, USA December 07-11, 2001 ISSN: 0006-4971.

L4 ANSWER 16 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. Zhang, Jiaoping (1); Chan, Edward K. L.; Lu, Maolong; Wang, Xiaoping; Xuanxian; Mueller-Pillasch, Friederike; Tan, Eng M. Autoimmune responses to mRNA binding proteins p62 and Koc in cancer. Proceedings of the American Association for Cancer Research Annual Meeting, (March, 2001) Vol. 42, pp. 157. print. Meeting Info.: 92nd Annual Meeting of the American Association for Cancer Research New Orleans, LA, USA March 24-28, 2001 ISSN: 0197-016X.

L4 ANSWER 17 OF 36 MEDLINE DUPLICATE 7 Zhang J Y; Chan E K; Peng X X; Lu M; Wang X; Mueller F; Tan E M Autoimmune responses to mRNA binding proteins p62 and Koc in diverse malignancies. CLINICAL IMMUNOLOGY, (2001 Aug) 100 (2) 149-56. Journal code: 100883537. ISSN: 1521-6616.

L4 ANSWER 18 OF 36 CAPLUS COPYRIGHT 2003 ACS Wang, Tongtong; Fan, Lijun Compounds and methods for therapy and diagnosis of lung cancer PCT Int. Appl., 261 pp. CODEN: PIXXD2

L4 ANSWER 19 OF 36 CAPLUS COPYRIGHT 2003 ACS DUPLICATE 8 Wang, Tongtong; Hopkins, Deborah; Schmidt, Cheryl; Silva, Sandra; Houghton, Raymond; Takita, Hiroshi; Repasky, Elizabeth; Reed, Steven G. Identification of genes differentially over-expressed in lung squamous cell carcinoma using combination of cDNA subtraction and microarray analysis. Oncogene (2000), 19(12), 1519-1528 CODEN: ONCNE5; ISSN: 0950-9232

L4 ANSWER 20 OF 36 MEDLINE DUPLICATE 9 Gure A O; Stockert E; Arden K C; Boyer A D; Viars C S; Scanlan M J; Old L J; Chen Y T CT10: a new cancer-testis (CT) antigen homologous to CT7 and the MAGE family, identified by representational-difference analysis. INTERNATIONAL JOURNAL OF CANCER, (2000 Mar 1) 85 (5) 726-32. Journal code: 0042124. ISSN: 0020-7136.

L4 ANSWER 21 OF 36 CAPLUS COPYRIGHT 2003 ACS Chen, Yao-Tsung; Gure, Ali; Tsang, Solam; Stockert, Elizabeth; Jager, Elke; Knuth, Alexander; Old, Lloyd J. Isolated nucleic acid molecule encoding the cancer associated antigens CT7, KOC-2 and KOC-3, expression of the antigens and application to cancer diagnosis and immunotherapy PCT Int. Appl., 45 pp. CODEN: PIXXD2

L4 ANSWER 22 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. Wallrapp, C.; Mueller-Pillasch, F.; Michha, A.; Wenger, C.; Geng, M.; Solinas-Toldo, S.; Lichter, P.; Frohme, M.; Hoheisel, J. D.; Adler, G.; Gress, T. M. (1) Strategies for the detection of disease genes in pancreatic cancer

- SO Lohr, J.-Matthias; Colcher, David; Hollingsworth, Michael A.; Liebe, Stefan. *Annals of the New York Academy of Sciences*, (June 30, 1999) Vol. 880, pp. 122-146. *Annals of the New York Academy of Sciences: Cell and molecular biology of pancreatic carcinoma: Recent developments in research and experimental therapy*. print.
- Publisher: New York Academy of Sciences 2 East 63rd Street, New York, NY, 10021, USA.
- Meeting Info.: Baltic Pancreas Meeting on Pancreatic Carcinoma from Bench to Bedside Warnemünde, Germany September 17-18, 1998
- ISSN: 0077-8923. ISBN: 1-57331-219-3 (cloth), 1-57331-220-7 (paper).
- ANSWER 23 OF 36 MEDLINE
Zhang J Y; Chan E K; Peng X X; Tan E M
DUPLICATE 10
- TI A novel cytoplasmic protein with RNA-binding motifs is an autoantigen in human hepatocellular carcinoma.
- SO JOURNAL OF EXPERIMENTAL MEDICINE, (1999 Apr 5) 189 (7) 1101-10. Journal code: 2985109R. ISSN: 0022-1007.
- L4 ANSWER 24 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
- AU Yamada, Akira (1); Kawano, Koichiro; Harashima, Nanae; Niya, Fumihiko; Nagai, Kouji; Kobayashi, Terutada; Mine, Takashi; Ushijima, Kimio; Nishida, Takashi; Itoh, Kyogo
- TI Study of HLA class I restriction and the directed antigens of cytotoxic T lymphocytes at the tumor sites of ovarian cancer.
- SO Cancer Immunology Immunotherapy, (May-June, 1999) Vol. 48, No. 2-3, pp. 147-152. ISSN: 0340-7004.
- L4 ANSWER 25 OF 36 MEDLINE
DUPLICATE 11
- AU Mueller-Pillasch F; Pohl B; Wilda M; Lacher U; Bell M; Wallrapp C; Hamelster H; Knochel W; Adler G; Gress T M
- TI Expression of the highly conserved RNA binding protein KOC in embryogenesis.
- SO MECHANISMS OF DEVELOPMENT, (1999 Oct) 88 (1) 95-9. Journal code: 9101218. ISSN: 0925-4773.
- L4 ANSWER 26 OF 36 CAPLUS COPYRIGHT 2003 ACS
- AU Chen, Yao-Tseng; Gure, Ali O.; Tsang, Solam; Stockert, Elisabeth; Jager, Elke; Knuth, Alexander; Old, Lloyd J.
- TI Identification of multiple cancer/testis antigens by allogeneic antibody screening of a melanoma cell line library
- SO Proceedings of the National Academy of Sciences of the United States of America (1998), 95(12), 6919-6923
- CODEN: PNASAA6; ISSN: 0027-8424
- L4 ANSWER 27 OF 36 CAPLUS COPYRIGHT 2003 ACS
- AU Hirai, Nobuyuki
- TI Establishment and characterization of ovarian yolk sac tumor cell line in nude mice
- SO Kurume Igakkai Zasshi (1998), 61(12), 334-343
- CODEN: KIZAAI; ISSN: 0368-5810
- L4 ANSWER 28 OF 36 CAPLUS COPYRIGHT 2003 ACS
- AU Tomioka, Yoshito
- TI Establishment and characterization of three human ovarian clear cell carcinoma cell line
- SO Kurume Igakkai Zasshi (1998), 61(12), 323-333
- CODEN: KIZAAI; ISSN: 0368-5810
- L4 ANSWER 29 OF 36 MEDLINE
DUPLICATE 12
- AU Motomura S
- TI Induction of apoptosis in ovarian carcinoma cell line by gonadotropin-releasing hormone agonist.
- SO KURUME MEDICAL JOURNAL, (1998) 45 (1) 27-32.

Journal code: 2985210R. ISSN: 0023-5679.

L4 ANSWER 30 OF 36 MEDLINE
 AU Mueller-Pillasch F, Lacher U, Wallrapp C, Micha A, Zimmerhackl F, Hamelster H, Varga G, Fries H, Buchler M, Beger H G, Villa M R, Adler G, Gress T M
 TI Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein.
 SO ONCOGENE, (1997 Jun 5) 14 (22) 2729-33.
 Journal code: 8711562. ISSN: 0950-9232.
 L4 ANSWER 31 OF 36 CAPLUS COPYRIGHT 2003 ACS
 DUPLICATE 14
 AU Kataoka, Aki; Nishida, Takashi; Hirai, Nobuyuki; Tomioka, Yoshito; Sugiyama, Toru; Yakushiji, Michiaki
 TI Induction of apoptosis in ovarian carcinoma cell line by glucocorticoids, and sex steroid hormones
 SO Oncology Reports (1997), 4(6), 1249-1253
 CODEN: OCRPEW; ISSN: 1021-335X
 L4 ANSWER 32 OF 36 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
 AU Mueller-Pillasch, Friederike (1); Lacher, Ulrike (1); Wallrapp, Christine (1); Micha, Anne (1); Zimmerhackl, Frank (1); Varga, G.; Fries, H.; Buchler, M.; Villa, M. R.; Hamelster, H.; Adler, Guido (1); Gress, Thomas M. (1)
 TI Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein.
 SO Gastroenterology, (1997) Vol. 112, No. 4 SUPPL., pp. A621.
 Meeting Info.: Digestive Disease Week and the 97th Annual Meeting of the American Gastroenterological Association Washington, D.C., USA May 11-14, 1997
 ISSN: 0016-5085.
 L4 ANSWER 33 OF 36 SCISEARCH COPYRIGHT 2003 ISI (R)
 AU Okina H, Kataoka A (Reprint); Sugiyama T; Nishida T; Yakushiji M
 TI Establishment and characterization of ovarian endometrial adenocarcinoma cell line in nude mice and analyses of the immunohistochemical property among the original, recurrent, and heterotransplanted tumor
 SO INTERNATIONAL JOURNAL OF ONCOLOGY, (FEB 1997) Vol. 10, No. 2, pp. 311-316.
 Publisher: INT JOURNAL ONCOLOGY, C/O PROFESSOR D A SPANDIDOS, EDITORIAL OFFICE, 1, S MERKOURI ST, ATHENS 116 35, GREECE.
 ISSN: 1019-6439.
 L4 ANSWER 34 OF 36 MEDLINE
 DUPLICATE 15
 AU Kataoka A; Nishida T; Motomura S; Higashijima T; Ueyama T; Sugiyama T; Yakushiji M
 TI The effects of sex steroid hormones, gonadotropins, and gonadotropin releasing hormone agonist on the proliferation of ovarian cancer cell line (KOC-2S).
 SO NIPPON SANKA FUJINKA GAKKAI ZASSHI. ACTA OBSTETRICA ET GYNAECOLOGICA JAPONICA, (1994 Oct) 46 (10) 1033-40.
 Journal code: 7505749. ISSN: 0300-9165.
 L4 ANSWER 35 OF 36 CAPLUS COPYRIGHT 2003 ACS
 AU Akizuki, Hideaki
 TI An experimental study on optimal administration of cisplatin for ovarian cancer
 SO Kurume Igakkai Zasshi (1991), 54(5), 359-66
 CODEN: KIZAL; ISSN: 0368-5810
 L4 ANSWER 36 OF 36 MEDLINE
 DUPLICATE 16
 AU Kataoka A; Yokota D; Sugiyama T; Oda T; Nishida T; Yakushiji M
 TI Effects of anticancer agents on 7, 12-dimethylbenz (a) anthracene induced rat ovarian cancer cell line (DMBA-OC-1) and human ovarian serous adenocarcinoma cell line (KOC-1S).

SO NIPPON SANKA FUJINKA GAKKAI ZASSHI. ACTA OBSTETRICA ET GYNAECOLOGICA
JAPONICA, (1989 Oct) 41 (10) 1523-9.
Journal code: 7505749. ISSN: 0300-9165.

=> s (immune(w) response or t(w) cell) and l4

3 FILES SEARCHED...

L5 8 (IMMUNE(W) RESPONSE OR T(W) CELL) AND L4

=> d his

(FILE 'HOME' ENTERED AT 15:48:12 ON 24 APR 2003)

FILE 'MEDLINE', CAPLUS, BIOSIS, SCISEARCH' ENTERED AT 15:48:26 ON 24 APR
2003

L1 1253 S KOC
L2 0 S MUELLER-PALLASCH AND L1
L3 71 S L1 AND CANCER
L4 36 DUP REM L3 (35 DUPLICATES REMOVED)
L5 8 S (IMMUNE(W) RESPONSE OR T(W) CELL) AND L4

=> d au ti so ab 1-8 15

L5 ANSWER 1 OF 8 MEDLINE
AU Wang T; Fan L; Watanabe Y; McNeill P D; Moulton G G; Bangur C; Fanger G R;
TI L523S, an RNA-binding protein as a potential therapeutic target for lung
SO BRITISH JOURNAL OF CANCER, (2003 Mar 24) 88 (6) 887-94.
Journal code: 0370635. ISSN: 0007-0920.

AB Approaches to vaccine-based immunotherapy of human cancer may
ultimately require targets that are both tumour-specific and immunogenic.
In order to generate specific antitumour immune

responses to lung cancer, we have sought lung
cancer-specific proteins that can be targeted for adjuvant vaccine
therapy. By using a combination of cDNA subtraction and microarray
analysis, we previously reported the identification of an RNA-binding
protein within the KOC family, L523S, to be overexpressed in
squamous cell cancers of the lung. We show here that L523S
exhibits significant potential for vaccine immunotherapy of lung
cancer. As an oncofetal protein, L523S is normally expressed in
early embryonic tissues, yet it is re-expressed in a high percentage of
non-small cell lung carcinoma. The specificity of L523S expression in lung
cancer was demonstrated by both mRNA and protein measurements
using real-time PCR, Western blot, and immunohistochemistry analyses.
Furthermore, we show that immunological tolerance of L523S is naturally
broken in lung cancer patients, as evidenced by detectable
antibody responses to recombinant L523S protein in eight of 17 lung
pleural effusions from lung cancer patients. Collectively, our
studies suggest that L523S may be an important marker of malignant
progression in human lung cancer, and further suggest that
treatment approaches based on L523S as an immunogenic target are worthy of
pursuit. British Journal of Cancer (2003) 88, 887-894.
doi:10.1038/sj.bjc.6600806 www.bjcancer.com

L5 ANSWER 2 OF 8 MEDLINE
AU Zhang J Y; Chan E K; Peng X X; Lu M; Wang X; Mueller F; Tan E M
TI Autoimmune responses to mRNA binding proteins p62 and Koc in
SO CLINICAL IMMUNOLOGY, (2001 Aug) 100 (2) 149-56.
Journal code: 100883537. ISSN: 1521-6616.

AB Two tumor-associated antigens, p62 and Koc, are insulin-like
growth factor II (IGF-II) messenger RNA binding proteins. Autoantibodies
to p62 have been detected in cancer sera but have not been

reported for Koc. This study determined the extent and frequency of autoantibodies to p62 and Koc in diverse malignancies, the epitopes on the antigens, and the presence or absence of cross-reactive antibodies. Recombinant polypeptides were expressed from full-length and partial cDNA constructs and used as antigens in Western blotting, enzyme-linked immunosorbent assay, and immunoprecipitation. After identifying the epitopes, cross-absorption with recombinant polypeptides was used to determine specificity. Sera from 777 patients with 10 different types of malignancy were analyzed. Autoantibodies to p62 were found in 11.6% and to Koc in 12.2% and cumulatively to both antigens in 20.5%, with significant difference from the control populations consisting of normal subjects and autoimmune disease patients ($P > 0.01$). The immunodominant epitopes were at the amino terminus of both antigens and absorption studies showed that the majority of autoantibodies were not cross-reactive. Autoantibodies to p62 and Koc were present in approximately similar frequencies in a variety of malignancies and the **immune responses** appeared to be independent of each other. The **immune responses** might be related to overexpression or dysregulation of p62 and Koc in some tumors.

ANSWER 3 OF 8 MEDLINE
 Gure A O, Stockert E, Arden K C, Boyer A D, Viars C S, Scanlan M J, Old L J, Chen Y T
 CT10: a new cancer-testis (CT) antigen homologous to CT7 and the MAGE family, identified by representational-difference analysis. INTERNATIONAL JOURNAL OF CANCER, (2000 Mar 1) 85 (5) 726-32.
 Journal code: 0042124. ISSN: 0020-7136.
 Assays relying on humoral or T-cell-based recognition of tumor antigens to identify potential targets for immunotherapy have led to the discovery of a significant number of immunogenic gene products, including cancer-testis (CT) antigens predominantly expressed in cancer cells and male germ cells. The search for cancer-specific antigens has been extended via the technique of representational-difference analysis and SK-MBL-37, a melanoma cell line expressing a broad range of CT antigens. Using this approach, we have isolated CT antigen genes, genes over-expressed in cancer, e.g., PRAME and KOC, and genes encoding neuro-ectodermal markers. The identified CT antigen genes include the previously defined MAGE-A6, MAGE-A4a, MAGE-A10, CT7/MAGE-C1, as well as a novel gene designated CT10, which shows strong homology to CT7/MAGE-C1 both at cDNA and at genomic levels. Chromosome mapping localized CT10 to Xq27, in close proximity to CT7/MAGE-C1 and MAGE-A genes. CT10 mRNA is expressed in testis and in 20 to 30% of various human cancers. A serological survey identified 2 melanoma patients with anti-CT10 antibody, demonstrating the immunogenicity of CT10 in humans.

Copyright 2000 Wiley-Liss, Inc.

ANSWER 4 OF 8 MEDLINE
 Zhang J Y, Chan E K, Peng X X, Tan E M
 A novel cytoplasmic protein with RNA-binding motifs is an autoantigen in human hepatocellular carcinoma.
 JOURNAL OF EXPERIMENTAL MEDICINE, (1999 Apr 5) 189 (7) 1101-10.
 Journal code: 2985109R. ISSN: 0022-1007.

AB In hepatocellular carcinoma (HCC), autoantibodies to intracellular antigens are detected in 30-40% of patients. Patients with chronic hepatitis or liver cirrhosis develop HCC, and when this occurs, some patients exhibit autoantibodies of new specificities. It has been suggested that these novel autoantibody responses may be immune system reactions to proteins involved in transformation-associated cellular events. One HCC serum shown to contain antibodies to unidentified cellular antigens was used to immunoscreen a cDNA expression library, and a full length cDNA clone was isolated with an open reading frame encoding 556 amino acids with a predicted molecular mass of 62 kD. The 62-kD

protein contained two types of RNA-binding motifs, the consensus sequence RNA-binding domain (CS-RBD) and four hnRNP K homology (KH) domains. This protein, provisionally called p62, has close identity (66-70%) to three other proteins at the amino acid sequence level, and all four proteins may belong to a family having CS-RBD in the NH2-terminal region and four KH domains in the mid-to-COOH-terminal region. The homologous proteins are: KH domain-containing protein overexpressed in cancer (Koc); zipcode binding protein, a protein which binds to a conserved nucleotide element in chicken beta-actin mRNA (ZBP1); and a protein which binds to a promoter cis element in *Xenopus laevis* TFIID gene (B3). p62 protein is cytoplasmic in location, and autoantibodies were found in 21% of a cohort of HCC patients. Patients with chronic hepatitis and liver cirrhosis, conditions which are frequent precursors to HCC, were negative for these autoantibodies, suggesting that the immune response might be related to cellular events leading to transformation. However, the possible involvement of p62 autoantigen as a factor in the transformation process remains to be elucidated.

ANSWER 5 OF 8 CAPLUS COPYRIGHT 2003 ACS
IN Wang, Tongtong; Fan, Liqun
TI Compounds and methods for therapy and diagnosis of lung cancer
SO PCT Int. Appl., 261 pp.
CODEN: PIXXD2

AB Comps. and methods for the treatment and diagnosis of lung cancer
are provided. The inventive compds. include polypeptides contg. at least a portion of a lung tumor protein. Vaccines and pharmaceutical compns. for immunotherapy of lung cancer comprising such polypeptides, or DNA mols. encoding such polypeptides, are also provided, together with DNA mols. for prepq. the inventive polypeptides.

LS ANSWER 6 OF 8 CAPLUS COPYRIGHT 2003 ACS
IN Chen, Yao-Tseng; Gure, Ali; Tsang, Solam; Stockert, Elisabeth; Jager, Elke; Knuth, Alexander; Old, Lloyd J.
TI Isolated nucleic acid molecule encoding the cancer associated antigens CT7, KOC-2 and KOC-3, expression of the
SO PCT Int. Appl., 45 pp.
CODEN: PIXXD2

AB The invention relates to newly identified cancer assocd. antigens, referred to as CT7, KOC-2 and KOC-3. The invention also relates to observations regarding known mol. KOC-1. It has been discovered that each of these mols. provokes antibodies when expressed by a subject. The ramifications of this observation are also a part of this invention.

LS ANSWER 7 OF 8 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AU Tan, Eng M. (1)
TI Autoantibodies as reporters identifying aberrant cellular mechanisms in tumorigenesis.
SO Journal of Clinical Investigation, (November, 2001) Vol. 108, No. 10, pp. 1411-1415. print.
ISSN: 0021-9738.

LS ANSWER 8 OF 8 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
AU Yamada, Akira (1); Kawano, Koichiro; Harashima, Nanae; Niiya, Fumihiko; Nagai, Kouji; Kobayashi, Terutada; Mine, Takashi; Ushijima, Kimio; Nishida, Takashi; Itoh, Kyogo
TI Study of HLA class I restriction and the directed antigens of cytotoxic T lymphocytes at the tumor sites of ovarian cancer.
SO Cancer Immunology Immunotherapy, (May-June, 1999) Vol. 48, No. 2-3, pp. 147-152.
ISSN: 0340-7004.

AB The molecular basis of T-cell-mediated recognition of ovarian cancer cells remains to be fully addressed. In this

study we investigated HLA class I restriction and directed antigens of cytotoxic T lymphocytes (CTL) at the sites of ovarian cancer. Three HLA-class-I-restricted CTL lines were established from the tumor sites of ovarian cancer by culturing tumor-infiltrating lymphocytes or tumor-associated ascitic lymphocytes with interleukin-2: (1) HLA-A2402-restricted and ovarian-adenocarcinoma-specific CTL, (2) HLA-A2-restricted CTL recognizing histologically different cancers, and (3) HLA-B52-restricted and ovarian-cancer-specific CTL. HLA-A0201, HLA-A0206 and HLA-A0207 tumor cells were lysed by the HLA-A2-restricted CTL. HLA-B52 restriction of the third CTL line was confirmed by the transfection of HLA-B5201 cDNA into the tumor cells. The HLA-A2-restricted CTL recognized the SART-1, but not the MAGE-1 or MAGE-3 antigen. These results may facilitate a better understanding of the molecular basis of tumor-specific immunity at the tumor site of ovarian cancer.

=> d bib 5 6 15

ANSWER 5 OF 8 CAPLUS COPYRIGHT 2003 ACS

TI Compounds and methods for therapy and diagnosis of lung cancer
 DN 133:308983
 AN 2000:742121 CAPLUS
 IN Wang, Tongtong; Fan, Liqun
 PA Corixa Corporation, USA
 SO PCT Int. Appl., 261 pp.
 DT Patent
 LA English
 FAN,CNT 12

PATENT NO. KIND DATE APPLICATION NO. DATE

PI WO 2000061612 A2 20001019 WO 2000-US8896 20000403

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GU, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, ST, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, RM: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GN, GW, ML, MR, NE, SN, TD, TG

US 6482597 B1 20021119 US 2000-480884 20000110
 EP 1169347 A2 20020109 EP 2000-920102 20000403
 R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO

BR 2000009505 A 20020611 BR 2000-9505 20000403

JP 2002543769 T2 20021224 JP 2000-611554 20000403
 US 1999-285479 A 19990402
 US 1999-466396 A 19991217
 US 1999-476496 A 19991230
 US 2000-480884 A 20000110
 US 2000-510376 A 20000222
 US 1998-40802 A2 19980318
 US 1998-123912 A2 19980727
 US 1998-221107 A2 19981222
 WO 1999-US5798 W 19990317
 WO 2000-US8896 W 20000403

ANSWER 6 OF 8 CAPLUS COPYRIGHT 2003 ACS

AN 1999:691302 CAPLUS

TI Isolated nucleic acid molecule encoding the cancer associated

antigens CT7, KOC-2 and KOC-3, expression of the
antigens and application to cancer diagnosis and immunotherapy
Chen, Yao-Tseng; Gure, Ali; Tsang, Solam; Stockert, Elisabeth; Jager,
Eike; Knuth, Alexander; Old, Lloyd J.
Ludwig Institute for Cancer Research, USA
PCT Int. Appl., 45 pp.
CODEN: PIXXD2
DT Patent
LA English
FAN.CNT 1

PATENT NO. KIND DATE APPLICATION NO. DATE

PI WO 9954738 A1 19991028 WO 1999-US5766 19990316
W: AU, CA, CN, JP, KR, NZ, ZA
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL,
PT, SE

US 6297364 B1 20011002 US 1998-61709 19980417
CA 2324975 AA 19991028 CA 1999-2324975 19990316
AU 9930079 A1 19991108 AU 1999-30079 19990316
AU 745259 B2 20020314 EP 1999-911439 19990316
EP 1071957 A1 20010131 R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,
IE, FI

JP 2002512049 T2 20020423 JP 2000-545030 19990316
US 2002111470 A1 20020815 US 2001-899651 20010706
PRAI US 1998-61709 A 19980417
WO 1999-US5766 W 19990316
RE.CNT 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

=>

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 ; Search time 38 Seconds
(without alignments)
3139.507 Million cell updates/sec

Title: US-09-897-778-176

Perfect score: 2956

Sequence: 1 MKKLYIGNLSENAFSDLES.....VKHQOQKALQSGPPQSRKK 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organella.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.protist.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|--------|---------------------|
| 1 | 2956 | 100.0 | 579 | 4 | O00425 | Q00425 homo sapien |
| 2 | 2858 | 96.7 | 579 | 11 | O9CPN8 | Q3cpn8 mus musculus |
| 3 | 2483 | 84.0 | 593 | 13 | O57526 | Q57526 xenopus lae |
| 4 | 2461.5 | 83.3 | 594 | 13 | O73932 | O73932 xenopus lae |
| 5 | 2261.5 | 76.5 | 582 | 13 | O9P880 | Q9p880 brachydanio |
| 6 | 2219.5 | 75.1 | 576 | 13 | O42254 | O42254 gallus gall |
| 7 | 2192 | 74.2 | 577 | 11 | O88477 | O88477 mus musculus |
| 8 | 2181 | 73.8 | 577 | 4 | O9NZ18 | Q9nzi8 homo sapien |
| 9 | 1868 | 63.2 | 556 | 4 | O9Y6M1 | Q9y6m1 homo sapien |
| 10 | 835.5 | 28.3 | 558 | 5 | O3V269 | Q3v269 drosophila |
| 11 | 835 | 28.2 | 566 | 5 | O9NGS9 | Q9ngs9 drosophila |
| 12 | 473 | 16.0 | 100 | 11 | O9D054 | Q9d054 mus musculus |
| 13 | 418.5 | 14.2 | 854 | 5 | Q21605 | Q21605 caenorhabdi |
| 14 | 262.5 | 8.9 | 774 | 10 | O9L128 | Q9li28 oryza sativ |
| 15 | 260.5 | 8.8 | 542 | 10 | O8S7G1 | O8s7g1 oryza sativ |
| 16 | 242.5 | 8.2 | 557 | 5 | Q23487 | Q23487 caenorhabdi |

| | | | | | | |
|----|-------|-----|------|----|--------|---------------------|
| 17 | 240 | 8.1 | 644 | 4 | Q12828 | Q12828 homo sapien |
| 18 | 240 | 8.1 | 651 | 11 | Q91WJ8 | Q91wj8 mus musculus |
| 19 | 238 | 8.1 | 641 | 5 | O9BLA0 | Q9bla0 caenorhabdi |
| 20 | 236.5 | 8.0 | 568 | 10 | O9LXFS | Q9lxf5 arabidopsis |
| 21 | 235.5 | 8.0 | 621 | 10 | O9C553 | Q9c553 arabidopsis |
| 22 | 235.5 | 8.0 | 653 | 4 | O96AE4 | Q96ae4 homo sapien |
| 23 | 234 | 7.9 | 600 | 4 | O92946 | Q92946 homo sapien |
| 24 | 233 | 7.9 | 492 | 4 | O9UNW9 | Q9unw9 homo sapien |
| 25 | 233 | 7.9 | 498 | 4 | O43267 | O43267 homo sapien |
| 26 | 231.5 | 7.8 | 313 | 10 | O9LZ82 | Q9l282 arabidopsis |
| 27 | 231 | 7.8 | 364 | 5 | O95SZ9 | Q95sz9 drosophila |
| 28 | 230.5 | 7.8 | 493 | 11 | O9JKN6 | Q9jkn6 mus musculus |
| 29 | 229.5 | 7.8 | 618 | 5 | O9BLA1 | Q9bla1 caenorhabdi |
| 30 | 229.5 | 7.8 | 386 | 3 | O74919 | O74919 schizosacch |
| 31 | 229 | 7.7 | 386 | 5 | P91632 | P91632 drosophila |
| 32 | 228.5 | 7.7 | 313 | 4 | O96EP6 | Q96ep6 homo sapien |
| 33 | 226 | 7.6 | 1301 | 5 | O9U982 | Q9u982 drosophila |
| 34 | 226 | 7.6 | 1301 | 5 | O9V8H6 | Q9v8h6 drosophila |
| 35 | 225.5 | 7.6 | 577 | 10 | Q9SR13 | Q9sr13 arabidopsis |
| 36 | 225.5 | 7.6 | 589 | 5 | O17935 | Q17935 caenorhabdi |
| 37 | 225.5 | 7.6 | 611 | 5 | O17936 | Q17936 caenorhabdi |
| 38 | 224 | 7.6 | 769 | 13 | O8UVD9 | O8uvd9 gallus gall |
| 39 | 222 | 7.5 | 479 | 10 | O9X171 | Q9x171 arabidopsis |
| 40 | 219.5 | 7.4 | 396 | 13 | O9PS11 | Q9ps11 xenopus lae |
| 41 | 217.5 | 7.4 | 640 | 10 | O9ASX3 | Q9asx3 arabidopsis |
| 42 | 217.5 | 7.4 | 644 | 10 | O9FNK3 | Q9fnk3 arabidopsis |
| 43 | 215.5 | 7.3 | 403 | 11 | O9DB01 | Q9db01 mus musculus |
| 44 | 214 | 7.2 | 833 | 10 | O9FMF0 | Q9fmf0 arabidopsis |
| 45 | 213.5 | 7.2 | 510 | 10 | O8W5C2 | O8w5c2 oryza sativ |

ALIGNMENTS

| | | | | | |
|----------|---|--------------|------|---------|--|
| RESULT 1 | | | | | |
| O00425 | | PRELIMINARY; | PRT; | 579 AA. | |
| AC | O00425 | | | | |
| AT | 01-JUL-1997 (TRENBLrel. 04, Created) | | | | |
| DT | 01-JUN-2002 (TRENBLrel. 04, Last sequence update) | | | | |
| DT | 01-JUN-2002 (TRENBLrel. 21, Last annotation update) | | | | |
| DE | Putative RNA binding protein KOC (KOC). | | | | |
| GN | KOC. | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| OX | NCBI_TaxID=9606; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.; | | | | |
| RL | Oncogene 0:0-0(0). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE-PANCREAS; | | | | |
| RA | Mueller-Pillasch F., Lacher U., Wallrapp C.; | | | | |
| RL | Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases. | | | | |
| DR | EMBL; U97188; AAC35208.1; . | | | | |
| DR | EMBL; U76705; AAD09223.1; . | | | | |
| DR | InterPro; IPR004087; KH_dom. | | | | |
| DR | InterPro; IPR004088; KH_type_1. | | | | |
| DR | InterPro; IPR000504; RNA_rec_mot. | | | | |
| DR | Pfam; PF00013; KH-domain; 4. | | | | |
| DR | Pfam; PF00076; rrm; 2. | | | | |
| DR | SMART; SM00322; KH; 4. | | | | |
| DR | SMART; SM00360; RRM; 2. | | | | |
| DR | PROSITE; PS00084; KH_TYPE_1; 4. | | | | |
| DR | PROSITE; PS0102; RRM; 2. | | | | |
| DR | PROSITE; PS00030; RRM_RNP_1; FALSE_NEG. | | | | |
| SQ | SEQUENCE 579 AA; 63720 MW; AE5C3A8EE3C135C5 CRC64; | | | | |

Query Match 100.0%; Score 2956; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-187; Indels 0; Gaps 0;
Matches 579; Conservative 0; Mismatches 0;

QY 1 MNKLYIGNLSNAAPSDESIFPKDAKIPVSGPFLVKTGYAFVDCPDDESWALKAEIALSGK 60
 Db 1 MNKLYIGNLSNAAPSDESIFPKDAKIPVSGPFLVKTGYAFVDCPDDESWALKAEIALSGK 60
 QY 61 IELHGKPIEVSHSVKQRIRKLOIRNIPPHLOWEVLDSLLVQYGVESCEQVNTDSETA 120
 Db 61 IELHGKPIEVSHSVKQRIRKLOIRNIPPHLOWEVLDSLLVQYGVESCEQVNTDSETA 120
 QY 121 VVNVYSSKQDQARQALDKLNGFLENFTLKVAY IPDEMAAQNPLQOPRRGRLGQSGSS 180
 Db 121 VVNVYSSKQDQARQALDKLNGFLENFTLKVAY IPDEMAAQNPLQOPRRGRLGQSGSS 180
 QY 181 ROGSPGVSQKPCDLPRLVLTQVFGAILKEGATIRNITKQTSKIDVHRKENAGAA 240
 Db 181 ROGSPGVSQKPCDLPRLVLTQVFGAILKEGATIRNITKQTSKIDVHRKENAGAA 240
 QY 241 EKSTITLSTPEGTSAACKSILEIMHKEAQDIFKEEIPKLILAHNNFVGRIGKEGRNLK 300
 Db 241 EKSTITLSTPEGTSAACKSILEIMHKEAQDIFKEEIPKLILAHNNFVGRIGKEGRNLK 300
 QY 301 KIEQDQTKITISPLQELTYNPERITVKGNVETCAKAEEMKKIRESYENDIASMNL 360
 Db 301 KIEQDQTKITISPLQELTYNPERITVKGNVETCAKAEEMKKIRESYENDIASMNL 360
 QY 361 QAHLIPGLNLNALGLFPPTSGMPPTSGPPSAMTPPYQFQESSETETVHOFIPALSVGAI 420
 Db 361 QAHLIPGLNLNALGLFPPTSGMPPTSGPPSAMTPPYQFQESSETETVHOFIPALSVGAI 420
 QY 421 IKOGOHIKQLSRFAGASIKIAPAEADAKVRWVIITGPPPEAQKAGRIYGIKEENFV 480
 Db 421 IKOGOHIKQLSRFAGASIKIAPAEADAKVRWVIITGPPPEAQKAGRIYGIKEENFV 480
 QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNSAEVVPVPRDQTPDENDQVVKIT 540
 Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNSAEVVPVPRDQTPDENDQVVKIT 540
 QY 541 GHFYACQVQAKRIQELITQVQHQQKALQSGPPQSRRK 579
 Db 541 GHFYACQVQAKRIQELITQVQHQQKALQSGPPQSRRK 579

RESULT 2

Q9CPN8 PRELIMINARY: PRT: 579 AA.
 AC Q9CPN8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 10 days embryo cDNA, RIKEN full-length enriched library,
 DE clone:2610036B18, full insert sequence (19f2 mRNA-binding protein
 DE 3).
 GN IGP2BP3 OR 2610101N1RIK OR MIMP3.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staibli F., Suzuki R., Tomita N., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
 RA Hayashaw-Borls A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
 RA Yuasa Y., Takeda M., Okano H.;
 RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications
 RT for the developing central nervous system.";
 RL J. Neurosci. Res. 0:0-0(2001).
 DR EMBL; AK011689; BAB21779.1; -;
 DR MGI; AB046173; BAB19755.1; -;
 DR MGD; MGI:1890359; IGF2bp3.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR000113; KH-domain; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS00084; KH_type_1; 4.
 DR PROSITE; PS0102; RRM; 2.
 SQ SEQUENCE 579 AA: 63574 MW: CABD9A43558392B7 CRC64;

Query Match 96.7%; Score 2858; DB 11; Length 579;

Best Local Similarity 96.4%; Pred. No. 5.2e-181;

Matches 558; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNKLYIGNLSNAAPSDESIFPKDAKIPVSGPFLVKTGYAFVDCPDDESWALKAEIALSGK 60
 Db 1 MNKLYIGNLSNAAPSDESIFPKDAKIPVSGPFLVKTGYAFVDCPDDESWALKAEIALSGK 60
 QY 61 IELHGKPIEVSHSVKQRIRKLOIRNIPPHLOWEVLDSLLVQYGVESCEQVNTDSETA 120
 Db 61 IELHGKPIEVSHSVKQRIRKLOIRNIPPHLOWEVLDSLLVQYGVESCEQVNTDSETA 120
 QY 121 VVNVYSSKQDQARQALDKLNGFLENFTLKVAY IPDEMAAQNPLQOPRRGRLGQSGSS 180
 Db 121 VVNVYSSKQDQARQALDKLNGFLENFTLKVAY IPDEMAAQNPLQOPRRGRLGQSGSS 180
 QY 181 ROGSPGVSQKPCDLPRLVLTQVFGAILKEGATIRNITKQTSKIDVHRKENAGAA 240
 Db 181 ROGSPGVSQKPCDLPRLVLTQVFGAILKEGATIRNITKQTSKIDVHRKENAGAA 240
 QY 241 EKSTITLSTPEGTSAACKSILEIMHKEAQDIFKEEIPKLILAHNNFVGRIGKEGRNLK 300
 Db 241 EKSTITLSTPEGTSAACKSILEIMHKEAQDIFKEEIPKLILAHNNFVGRIGKEGRNLK 300
 QY 301 KIEQDQTKITISPLQELTYNPERITVKGNVETCAKAEEMKKIRESYENDIASMNL 360
 Db 301 KIEQDQTKITISPLQELTYNPERITVKGNVETCAKAEEMKKIRESYENDIASMNL 360
 QY 361 QAHLIPGLNLNALGLFPPTSGMPPTSGPPSAMTPPYQFQESSETETVHOFIPALSVGAI 420
 Db 361 QAHLIPGLNLNALGLFPPTSGMPPTSGPPSAMTPPYQFQESSETETVHOFIPALSVGAI 420
 QY 421 IKOGOHIKQLSRFAGASIKIAPAEADAKVRWVIITGPPPEAQKAGRIYGIKEENFV 480
 Db 421 IKOGOHIKQLSRFAGASIKIAPAEADAKVRWVIITGPPPEAQKAGRIYGIKEENFV 480
 QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNSAEVVPVPRDQTPDENDQVVKIT 540
 Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNSAEVVPVPRDQTPDENDQVVKIT 540
 QY 541 GHFYACQVQAKRIQELITQVQHQQKALQSGPPQSRRK 579
 Db 541 GHFYACQVQAKRIQELITQVQHQQKALQSGPPQSRRK 579

RESULT 3
 ID 057526 PRELIMINARY; PRT; 593 AA.
 AC 057526;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE KH domain-containing transcription factor B3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92249652; PubMed=1577195;
 RA Pfaff S.L., Taylor W.L.;
 RT "Characterization of a Xenopus oocyte factor that binds to a
 RT developmentally regulated cis-element in the TFIIB gene."
 RL Dev. Biol. 151:306-316(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX Griffin D., Taylor W.L.;
 RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
 RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
 RL Genes Dev. 0:0-0(1998).
 DR EMBL; AF042353; AAB97457.1; -;
 DR EMBL; AF046633; AAC18597.1; -;
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH-domain; 4.
 DR Pfam; PF00076; rim; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 SQ SEQUENCE 593 AA; 65385 MW; 5A5ABABAAID55DF7 CRC64;
 Query Match 84.0%; Score 2483; DB 13; Length 593;
 Best Local Similarity 82.1%; Pred. No. 3.6e-156;
 Matches 487; Conservative 47; Mismatches 45; Indels 14; Gaps 6;
 QY 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGFLVKTGYAFVDCPDSESWALKAIKALSGK 60
 DB 1 MNKLYIGNLSENVSPDLESIPKESKIPPTGQFLVKSQYAFVDCPDSEWALKAIKALSGK 60
 QY 61 IELHGKPIEVEHSVPKQRIRKIQIRNIPHLQNEVLDSLVLQGVVCEQVNTDSETA 120
 DB 61 VELHGKPIEVEHSVPKQRIRKIQIRNIPHLQNEVLDSLVLQGVVCEQVNTDSETA 120
 QY 121 VNVVYSSKDQARQALDKLNGFLENFTLKVAYIPDEMAAQNPLOQ-----PRGR 172
 DB 121 VNVVYANKERHARQLEKNGYOLENYSKVITYIPDEMATQPSQOOLQPOQHPOGR 180
 QY 173 ELGGRSSRQSPGSGVSKQRP-CDLPLRLVLPVTPQVGAIGKEGATIRNITKOTOSKIDV 231
 DB 181 SFGQGPARGQSPGAARPKQSEVPLRLVLPVTPQVGAIGKEGATIRNITKOTOSKIDI 240
 QY 232 HKENGAARKSITLSTPSTGSAACKSILETMHKEAODKFTETIPLKILAHNFPVGLR 291
 DB 241 HKENGAARKSPITHTSTPGCSAAKCIIMETMQEAODKFTETIPLKILAHNFPVGLR 300
 QY 292 LGKEGRNLKKIEQDQTDKITISPLQELTLNPERITIVKGNVETCAKAEIEIMKKIRTSY 351
 DB 301 LGKEGRNLKKIEQDQTDKITISPLQELTLNPERITIVKGSIEFCAKAESEVMKKIRTSY 360
 QY 352 ENDIAAMNLOAHLIPGLNLAALGLFP-PSGMPPTSPGSPAMT-PPYDF-EOSTETV 408

```

QY 231 VHRKENAGAAEKITITLSTPEGTSAAKSIILEIMHKEAQDIKTEETIPLKILAHNNFVR 290
DB 241 IHRKENAGAAEKITITLSTPEGTSAAKSIILEIMHKEAQDIKTEETIPLKILAHNNFVR 300
QY 291 LIGKEGNLAKIEQDITITLSTPEGTSAAKSIILEIMHKEAQDIKTEETIPLKILAHNNFVR 350
DB 301 LIGKEGNLAKIEQDITITLSTPEGTSAAKSIILEIMHKEAQDIKTEETIPLKILAHNNFVR 360
QY 351 YENDIASNNLQALHIFGLNLNALGLFP--PTSGMPPPTSGLPPS--AMTPPYPOF--EQSETET 407
DB 361 YENDIASNNLQALHIFGLNLNALGLFPSSSGMPPPTSGLPPS--AMTPPYPOF--EQSETET 420
QY 408 VHQFIPALSGVAGIIGKQGHQIKQLSRFAGASIKIAPADAKVRMVIITGPPPEAQFKAQ 467
DB 421 VHLFIPALSGVAGIIGKQGHQIKQLSRFAGASIKIAPADAKVRMVIITGPPPEAQFKAQ 480
QY 468 GRIYGTKEENFYSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVPDQ 527
DB 481 GRIYGTKEENFYSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVPDQ 540
QY 528 TPDENDQVVKITGHFYACQVAORKIQEILTVK--QHQQOKALQSGPPQSRK 579
DB 541 TPDENDQVVKITGHFYACQVAORKIQEILTVK--QHQQOKALQSGPPQSRK 594

RESULT 5
QSPW80 PRELIMINARY; PRT; 582 AA.
AC QSPW80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Vg1 RNA binding protein.
GN DVIRBP.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
RA Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;
FT "Vg1 RBP intracellular distribution and evolutionarily conserved
RT expression suggest multiple roles during development.";
RL Mech. Dev. 0:0-0(1999).
DR EMBL; AF161270; AAD45610.1; -.
DR ZFIN; ZDB-GENE-000308-1; dvirrbp.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH-domain; 4.
DR Pfam; PF00076; rtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_type_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 582 AA; 63351 MW; 9DAE63200681B306 CRC64;

Query Match 76.5%; Score 2261.5; DB 13; Length 582;
Best Local Similarity 77.1%; Pred. No. 1.6e-141;
Matches 450; Conservative 50; Mismatches 77; Indels 7; Gaps 5;

QY 1 MNKLYIGNLSENAPSDLESIFDKAPIVSGPFLVKTGYAFVDCPDSESWALKATELSGK 60
DB 1 MNKLYIGNVSEQASALDLESIFEQWKIPFSAPFLVKSGYAFVDCPDSESWALKATELSGK 60
QY 61 IELHGVPIEVHSVPRKQIRKLOIRNIPPHLOWELVLDLSLYQGVVSCQVNTDSETA 120
DB 61 VELHGVPIEVHSVPRKQIRKLOIRNIPPHLOWELVLDLSLYQGVVSCQVNTDSETA 120
QY 121 VVNVYSSKQQAQALDKINGQLENFTLVKVIIPDEMAAQNPLOQPRGRRLGQRGSS 180

```

```

DB 121 VVNVYSGAKDQAREAMDKLNGFLMENYALKVSYIPDETAADAP--AVGRRGFNRCPP 178
QY 181 ROGSPGYSKOK--PCDLPRLVPTQFVCAIIGKCATIRNITKOTQSKTDVHRKENAGA 239
DB 179 ROGSPGLARPKLQSDVPLRLVPTQFVCAIIGKCATIRNITKOTQSKTDVHRKENAGA 238
QY 240 AKSITLSTPEGTSAAKSIILEIMHKEAQDIKTEETIPLKILAHNNFVRIGKEGNL 299
DB 239 AKSITLSTPEGTSAAKSIILEIMHKEAQDIKTEETIPLKILAHNNFVRIGKEGNL 298
QY 300 KKEQDQDTKITISPLQELTYLPNTITVKNVETCAKAEETIMKKIRESYENDIASNN 359
DB 299 KKEQDQDTKITISPLQELTYLPNTITVKNVETCAKAEETIMKKIRESYENDIASNN 358
QY 360 LOAHILPGLNLNALGLFP--PTSG--MPPPTSGLPPS--AMTPPYPOF--EQSETETVHOFIPAL 415
DB 359 LOAHILPGLNLNALGLFP--PTSG--MPPPTSGLPPS--AMTPPYPOF--EQSETETVHOFIPAL 418
QY 416 SVGAIIGKQGHQIKQLSRFAGASIKIAPADAKVRMVIITGPPPEAQFKAQRIYGTIK 475
DB 419 AVGAIIGKQGHQIKQLSRFAGASIKIAPADAKVRMVIITGPPPEAQFKAQRIYGTIK 478
QY 476 EENFYSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVPDQTPDENDQV 535
DB 479 EENFYSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVPDQTPDENDQV 538
QY 536 VVKITGHFYACQVAORKIQEILTVKQHQQOKALQSGPPQSRK 579
DB 539 VVKITGHFYACQVAORKIQEILTVKQHQQOKALQSGPPQSRK 582

RESULT 6
O42254 PRELIMINARY; PRT; 576 AA.
AC O42254;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE "Characterization of a beta-actin mRNA zipcode-binding protein.";
DR Mol. Cell. Biol. 17:2158-2165(1997).
DR EMBL; AF026527; AAB82295.1; -.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH-domain; 4.
DR Pfam; PF00076; rtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_type_1; 4.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AAEF2D1D81C8811 CRC64;

Query Match 75.1%; Score 2219.5; DB 13; Length 576;
Best Local Similarity 74.7%; Pred. No. 9.4e-139;
Matches 437; Conservative 61; Mismatches 72; Indels 15; Gaps 7;

QY 1 MNKLYIGNLSENAPSDLESIFDKAPIVSGPFLVKTGYAFVDCPDSESWALKATELSGK 60
DB 1 MNKLYIGNLSENAPSDLESIFDKAPIVSGPFLVKTGYAFVDCPDSESWALKATELSGK 60

```


QY 61 IELHGKPIEVHSVPKQRIKRLQIRNIPPHLOWEVLDSLLVQYGVVESCQVNTDSETA 120
 DB 61 VELHGKQLEIHSVPKQKRSKTIQIRNIPQLRWEVLDCLLAQYGVVESCQVNTDSETA 120
 QY 121 VNVVYSSKDOARQALDKNGFLENFTLVKAYIPDEMAAQNPLQOPGRRR-GLGQGRS 179
 DB 121 VNVVYTNREQTROAIMKLNHQLNHLVKSYPDEQSVQ----GPENRRGGFGARGA 176
 QY 180 SRQGSP---GSVSKOKPCDPLRLVPTQVGNALIGKEGATIRNITKQTSKIDVHRKEN 236
 DB 177 PROGSPVTAAGVAKQOPVDIPRLVPTQVGAIGKEGATIRNITKQTSKIDVHRKEN 236
 QY 237 AGAAEKSTILSTPEGTSACKSILEIMHKEAQDIFTEIPLKTLAHNNFVGLIGKEG 296
 DB 237 AGAAEKASISHTPBGCSAACKMILEIMQKEAKDKTADDEVPLKTLAHNNFVGLIGKEG 296
 QY 297 RNLKKTQDTPKIFISPLQELTYNPERITITVKGNVETCAKAEIEIMKKITRESYENDIA 356
 DB 297 RNLKKEVDQETKIFISPLQELTYNPERITITVKGNSIENCKAEOEIMKKVREAYENDVA 356
 QY 357 SMNLQHLIPGLNMLGLPPTSGMPPPTSGPPSAMT---PPYPOFQESQETVHQFIPA 414
 DB 357 AMSLOSHLIPGLNLAAGLFPASSNAVPP---PPSSVSGAAPPYSPFPEQETVHVFIIPA 413
 QY 415 LSVGALIGKQGHIKOLSRFAGASIKIAPAEAPDAKVRVVIITGPPEAQKAGRIYCKI 474
 DB 414 QAVGALIGKQGHIKOLSRFAGASIKIAPPETPDSKVRVVIITGPPEAQKAGRIYCKL 473
 QY 475 KEENFVSPKEEVKLEAHLRPSFAAGRVIGKGGKTYNELONLSSAEVVPDPDQDENQ 534
 DB 474 KEENFPGKEEVKLETHLRVPSAAGRVIGKGGKTYNELONLTAEEVVPDPDQDENQ 533
 QY 535 VVYKIGTHPYACQVAKRIQBIITQVKKOHOQKALQSGPPQSRKK 579
 DB 534 VVYKIGTHPYASQAKRIQDILAQVKQ-QHQRG-QSGQLQARRK 576

RESULT 7
 O88477 PRELIMINARY; PRT; 577 AA.
 ID O88477; PRELIMINARY; PRT; 577 AA.
 AC O88477; 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Coding region determinant binding protein (Coding region determinant-binding protein).
 DE binding protein).
 GN IGF2BP1 OR CRDBP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92217743; PubMed=1559612;
 RA Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;
 RT "Control of c-myc mRNA half-life in vitro by a protein capable of binding to a coding region stability determinant.";
 RL Genes Dev. 6:642-654(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=941158886; PubMed=8114742;
 RA Herrick D.J., Ross J.;
 RT "The half-life of c-myc mRNA in growing and serum-stimulated cells: influence of the coding and 3' untranslated regions and role of ribosome translocation.";
 RL Mol. Cell. Biol. 14:2119-2128(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94179348; PubMed=8132663;
 RA Prokipcak R.D., Herrick D.J., Ross J.;
 RT "Purification and properties of a protein that binds to the C-terminal coding region of human c-myc mRNA.";
 RL J. Biol. Chem. 269:9261-9269(1994).

RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97322234; PubMed=9178888;
 RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.;
 RT "Developmental regulation of CRD-BP, an RNA-binding protein that stabilizes c-myc mRNA in vitro.";
 RL Oncogene 14:1279-1286(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A., Fleisig A.J.;
 RA Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Cojocori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AF061569; AAC72743.1; -;
 DR EMBL; AK013940; BAB29071.1; -;
 DR HSSP; P11940; 1CVJ
 DR MGD; MGI:1890357; Igf2bp1.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH-domain; 4.
 DR Pfam; PF00076; rim; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50084; KH_type_1; 4.
 DR PROSITE; PS50102; RRM; 2.
 SQ SEQUENCE 577 AA; 63450 MW; EFBBLAF2FF9F0344 CRC64;

Query Match 74.2%; Score 2192; DB 11; Length 577;
 Best Local Similarity 74.1%; Pred.No. 6.2e-137;
 Matches 434; Conservative 62; Mismatches 17; Indels 16; Gaps 8;

QY 1 MNKLYIGNLSSNAAPSDFIESFKADIKIPVSGPLVKTGYAFVDCPDSEWALKAEALSCK 60
 DB 1 MNKLYIGNLSSNAAPSDFIESFKADIKIPVSGPLVKTGYAFVDCPDSEWALKAEALSCK 60
 QY 61 IELHGKPIEVHSVPKQRIKRLQIRNIPPHLOWEVLDSLLVQYGVVESCQVNTDSETA 120
 DB 61 VELHGKQLEIHSVPKQKRSKTIQIRNIPQLRWEVLDCLLAQYGVVESCQVNTDSETA 120
 QY 121 VNVVYSSKDOARQALDKNGFLENFTLVKAYIPDEMAAQNPLQOPGRRR-GLGQGRS 179
 DB 121 VNVVYTNREQTROAIMKLNHQLNHLVKSYPDEQSVQ----GPENRRGGFGARGA 176
 QY 180 SRQGSP---GSVSKOKPCDPLRLVPTQVGNALIGKEGATIRNITKQTSKIDVHRKEN 236
 DB 177 PROGSPVTAAGVAKQOPVDIPRLVPTQVGAIGKEGATIRNITKQTSKIDVHRKEN 236
 QY 237 AGAAEKSTILSTPEGTSACKSILEIMHKEAQDIFTEIPLKTLAHNNFVGLIGKEG 296

QY 121 VVNTYSSKQDQARQALDKLNGFLENFTLKVAYIPDEMAAQNPLQOPRRGRGLGQSS 180
 DB 122 VVNTYATREAAKIAEKLGHQFENYSFKISYIPDEEVSSPSPQ--RAQR--GDHSSR 177
 QY 181 RQG-SPGSVKOKPCDPLRLVPTQFVCAIIGKSGATIRNITKQTSKIDVHKEKAGA 239
 DB 178 EOGHAPGTSOARQIDFPLRLVPTQFVCAIIGKSGATIRNITKQTSKIDVHKEKAGA 237
 QY 240 AKSITLSTPESTSAACKSILEIMHKEAQDIKFTTEIPLKILAHNNFVGRILGKEGNL 299
 DB 238 AKPVTTHATPEGSTSEACRMILSMOKEADETKLAETPLKILAHNNFVGRILGKEGNL 297
 QY 300 KKIQDQDTTKITSPLOELTYNPTITVKNVETCAKAEETMKKIKRESYENDIASMN 359
 DB 298 KRIEHTGKTITSSQDLSTIYPTITVKNVETCAKAEETMKKIKRESYENDIASMN 357
 QY 360 LOAHLIPGLNINLALGFPPTSGMPPPTSGPSPAMTPPYQFEQ-SETETVHOFTIPALSVG 418
 DB 358 THS-----GVF-----SSLYPHHQFGFPFHHSYQEQIVNLFPTQAVG 397
 QY 419 ALIIGKOGHOKOLSRFAGASIKITAPAEAPDAKRVMIITGPPEAQFKAQGRIGYKIKKEN 478
 DB 398 ALIIGKGAHILKQARFAGASIKITAPAEAPDAKRVMIITGPPEAQFKAQGRIGYKIKKEN 457
 QY 479 FVSPKEVKELAHTRVPSFAGRVIGKGGKTVNELONLSAEEVVVPRDQTPDNDQVVVK 538
 DB 458 FNPKEEVKELAHTRVPSFAGRVIGKGGKTVNELONLSAEEVVVPRDQTPDNDQVVVK 517
 QY 539 ITGHFYACQVQAQRIQIQLTVKQHQHQ 566
 DB 518 IIGHFASQTAQRIQIQLTVKQHQHQ 545

RESULT 10
 Q9VZ69 ID Q9VZ69 PRELIMINARY; PRT; 558 AA.
 AC Q9VZ69;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE CG1691 protein.
 GN IMP OR CG1691.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champs W., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler K.J., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Heston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Spirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003484; AAP47958.1; -.
 DR FlyBase; FBgn0030235; Imp.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH-domain; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PS00084; KH_TYPE_1; 4.
 SQ SEQUENCE 558 AA; 61748 MW; 833C9C56954C139C CRC64;

Query Match 28.3%; Score 835.5; DB 5; Length 558;
 Best Local Similarity 38.5%; Pred. No. 3.6e-47;
 Matches 198; Conservative 86; Mismatches 147; Indels 83; Gaps 12;

QY 121 VVNTYSSKQDQARQALDK-----LNGFLENFTLKVAYIPDEMAAQNPLQOPRRGRGLGQ 176
 DB 14 ISNNYQKQSLIRYLDRAAVGLNGVEFEG-----SKLHAEQLDKNRRSQ----- 60

QY 177 RGSROGSPGSVSKOKPCDPLRLVPTQFVCAIIGKSGATIRNITKQTSKIDVHKEK 236
 DB 61 --NORNPYPGMPGPGROADFPRLIVQSEVGAIIIGROGSTIRITQOSRARVDVHKEK 118

QY 237 AGAAEKSTITLSTPEGSTSAACKSILEIMHKEAQDIKFTTEIPLKILAHNNFVGR 290
 DB 119 VGSLEKSITITYGPNENCNTNACKRILEVMQOEAIKSTNKGELSPECSEICLKILAHNNLIGR 178

QY 291 LIKGEKNLAKKIEQDQDTKITISPLQELTYNPTITVKNVETCAKAEETMKKIKRES 350
 DB 179 IIGKSGNTIKRIMQDQDTKITVSSINDINSFLERITVKGLEIENMSRAEQISTKLKRS 238

QY 351 YENDIASNINLQAHILPGLNINLALGFPPTSGMPPPTSGP-----PSAMTPPY 398
 DB 239 YENDLQAMAPQSLMFPGLHPNMA-MSTPGNGVMVNTSMPPFSCQSFAMSKTPASVPPV- 296

QY 399 OFEQSETETVHOFTIPALSVGAIIIGKOGHOKOLSRFAGASIKITAPAE-----PDAKVRMV 454
 DB 297 -FPNDLQETTVLYIPNNAVGAIIIGTRGSHIRSMRFSNLSKILAPLDADKPLDQOTERKV 355

QY 455 IITGPPPEAQFKAQRIQIQLTVKQHQHQ-----DTPDENQVVKITGHEYACQVQAQRIQIQL- 514
 DB 356 TIVGTPGEGWKAQRIQIQLTVKQHQHQ-----DTPDENQVVKITGHEYACQVQAQRIQIQL- 415

QY 515 NLSSAEVVVPR-----DTPDENQVVKITGHEYACQVQAQRIQIQL----- 557
 DB 416 RVTGSVILKPEHALAPPSGGDEETP-----VHIIGLFYSVQSAQRIRRAMLSTNPPPP 468

QY 558 -----TQVKHQHQQA-----LQSGPPQS 576
 DB 469 ITKKQAKAEQLQQQQQSLAGASSGQQQQPQS 502

RESULT 11
 Q9NGS9 ID Q9NGS9 PRELIMINARY; PRT; 566 AA.
 AC Q9NGS9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
GN IGF-II mRNA-binding protein (SD07045p).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephyrgota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyrgoidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RT "Cloning and Expression of a Drosophila Homolog of IMP/Vg1-RBP."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STAPLETON M., BROCKSTEIN P., HONG L., AGBAYANI A., CARLSON J.,
RA CHAMPE M., CHAVEZ C., DORSETT V., TARTAN D., FRISSE E., GEORGE R.,
RA GONZALEZ M., GUARIN H., LI P., LIAO G., MIRANDA A., MANGALL C.J.,
RA NUNOO J., PACLEB J., PARAGAS V., PARK S., PHOUANENAVONG S., WAN K.,
RA YU C., LEWIS S.E., RUBIN G.M., CEJNLIK S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF241237; AAF63331.1;
DR EMBL; AY069821; AAL39966.1;
DR FlyBase; FBgn0030235; Imp.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH-domain; 4.
DR SMART; SM00322; KH; 4.
DR PROSITE; PS50084; KH_TYPE_1; 4.
SQ SEQUENCE 566 AA; 62129 MW; B4C74C5C99B0C030 CRC64;

Query Match 28.2%; Score 835; DB 5; Length 566;
Best Local Similarity 38.8%; Pred. No. 3.9e-47;
Matches 197; Conservative 86; Mismatches 147; Indels 78; Gaps 12;

Qy 121 VNVTVSSKDDARQALDK----LNGFQLENFTLKVAYIPDMAAQNPLQQPRGRGLGQ 176
Db 14 ISNNYQQKSLRYLDRAVNLGVGEFEG-----SKLHAEQLDKNQRRSQR---- 60

Qy 177 RGSRSRGSFGSVSKQPCDPLRLVLTQFVGAIGKEGATRNITKQSKIDVHRKEN 236
Db 61 --NORNPYGPMPGQGRQADFLRLVLQSEMVGALIGQSGTIRITQQSRAVDVHRKEN 118

Qy 237 AGAAKSITILSTPECTSAACKSILEIMHKEADQIKFTEIPLKILAHNNFVGLRIGKEG 296
Db 119 VGSLEKSITTYGNPENCTNACKRILEVMQOEASTN-KGEICKILAHNNLIGRIKSG 177

Qy 297 RNLKKTEDDTKTIISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIREYENDIA 356
Db 178 NTKRIMQDTRKITVSSINDINGFNLERIITVKGLENNRAENQISTKLQSYENDLQ 237

Qy 357 SMNLQALIPGLNALGLFPPTSCMPPTSGP-----PSAMTPPYPOEQSE 404
Db 238 AMAPOSLMFPGLHPWAM--MSTPGNGVFNFTSMPPFQSFAMSKTPASVVPEV--FPNDL 294

Qy 405 TETVHOFIPALSGVGAIGKQGHITKOLSRFAGASIKITAPAEA---PDAKVRMWLTGPP 460
Db 295 QETTYLYIPNNAVGALITGRSHIRSMRFSNASLAIAPLDADKPLDQOTKRTVITGTP 354

Qy 461 EAQFKAQRIYKIKTEENFVSKVEKLEAHIRVPSFAGRVIGGKGTNVLQNLSSAE 520
Db 355 EGQWKAQYWFIEKMEEGFMCGTDDRLTVELLVASSQVGRIGKGGQVRELQRTVGSV 414

Qy 521 VVYPR-----DOTPDENDQVVKITGHEFVACQVAKRKTQEL----- 557
Db 415 IKLPEHALAPPSGGDEETP-----VHIIGLFYSVQSAQRIRRAMLSTNPPPTTKQK 467

Qy 558 ---TQVQHQOQKA-----LQSGPPQS 576
Db 468 AAKEQLQQQQQSLAGASSGQQOQOOPS 495

RESULT 12

```

```

Q9D054
ID Q9D054 PRELIMINARY; PRT; 100 AA.
AC Q9D054;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2610101NIRIK protein.
GN IGF2BP3 OR 2610101NIRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK011797; BAB27848.1;
DR MGI; MGI:1890359; I9f2dp3.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 100 AA; 11249 MW; 4D871E37EB9D5466 CRC64;

Query Match 16.0%; Score 473; DB 11; Length 100;
Best Local Similarity 88.7%; Pred. No. 3.1e-24;
Matches 86; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNLSENAPSDLESIFKDAKIPVSGPFLVKTGTAFAVDCPDSEWALKAEALSGK 60
Db 1 MNKLYIGNLSDHAGPADLESFVKDAKIPVAGPFLVKTGTAFAVDCPDSEWALKAEALSGK 60

Qy 61 TELHGKPIEVHSVPKQRIRKLRIRKLRIRKLRIRKLRIRKLRIRKLRIRKLRIRKLRIRK 97
Db 61 MELHGKPMVEHSVPKQRIRKLRIRKLRIRKLRIRKLRIRKLRIRKLRIRKLRIRK 97

RESULT 13
Q21605
ID Q21605 PRELIMINARY; PRT; 854 AA.
AC Q21605;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE M88.5 protein.
GN M88.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;

```

RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99069613; PubMed=9851916;
 RX none;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z34802; CAA84338.1; .
 DR InterPro; IPR004087; KH_dom.
 DR Pfam; PF00013; KH_domain; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 SQ SEQUENCE 854 AA; 94553 MW; D3B40412353ECC21 CRC64;

Query Match 14.2%; Score 418.5; DB 5; Length 854;
 Best Local Similarity 24.1%; Pred. No. 2.6e-19;
 Matches 160; Conservative 103; Mismatches 215; Indels 187; Gaps 27;
 QY 78 QIRKLRINPPHQLQWEVLDSLLVQGVVSCQVNTDSETAVVN-----VTYSSK 129
 Db 167 QOMQOQOQNSGSAHFHQQLQAVQOQ-----QAQMHHRLQAGAPINQOQFVPPPTMWP 221
 QY 130 DOARALDKLNGFQLENTLKVAVIPDEMA--AQ--NPLQPRGRRLGQSGSSROGSP 185
 Db 222 QOMQQA-QOQQAQOQHOHQIHHQHPQMMQOHAQGVYHPOQNOQHOA-QOQOQSHHQSQ 279
 QY 186 G-----SVSKQKP-----C--DLRLRLVPTQFVGAILKKGATIRNTKQ 224
 Db 280 NNHQRHNQSHSGPHHTPQNLMPRCMLKDWPIRCVVEGYHAVIGPNSGTIKDIASS 339
 QY 225 TQSKID----VIRKENA--GAEEKSTITLSTPEGTSAACKSILEIMHKEA--QDIKFTEEI 277
 Db 340 TRCRVDFVNLKSKERTVLGNDRILTVHGAEQATKAVARILDVIOSEAVKDDVNVGADT 399
 QY 278 PLKILAHNFVRLIGKEGRNLKIEQDTDFKITISPLQELTLYNPRTITVKN 332
 Db 400 VLRRAHQNLGRLIGKAGSSIKELIMOKTGNITVTYIEPPGGISGLTANELLGLMERT 459
 QY 327 ITVKG-NYETCAKEEIMKKIRESYENDIASMNLQALHPLGLNLNALGLFPPTSGMPP- 384
 Db 460 LMRGPSIEAVVQAEALISAKLKCYESD-SOLRAQSNQCP-----MPPMK-MPPI 508
 QY 385 -PTSGPPSAMTPPY-----PQFQSETETVHQ----- 410
 Db 509 LPPGASSAVSAPHFIPTPVGMQVQHFASSQHLVHONANNSFLQPGVLQIQPGTTNLRQ 568
 QY 411 ---FTPALSVGAILKGQGHQIKQLSRFAGASIKIAPAEAPDAKV----- 451
 Db 569 VRMVVPSDMIGALIGAKGNKIMIRDTGASVKI---EAEPEKTQREAEAEKKRKLDET 625
 QY 452 -----RMVITGPPEAQFKAQGRIGYK 473
 Db 626 DSGCVASGDHPQEFLEDNATINSSDAIEKPKVSRVMTINGDDLQLLKQASYFSK 685
 QY 474 IKEENFVSP-----KEEVKLEAHTVPSFAGRVIGGKTVNELNLSAEVVPVR 525
 Db 686 IAEYSSSLPSGMDGRSHMLRIKTEVSVPTRIIGRIIGKGGQNVRELQITGAVVKIPE 745
 QY 526 DQTPD-----ENDQVVKITGHFYACQVAKRIQEIITGVQKQ---HQOQKA-LOS 571
 Db 746 EERNQGEVYRHDDGLEEDMTNIRTIGNWYTHNVQFRLAHVNEYRSGDHRNKSDDYK 805
 QY 572 GPPOS 576
 Db 806 GRPHS 810

RESULT 14
 Q9LI28 PRELIMINARY; PRT; 774 AA.
 ID Q9LI28
 AC Q9LI28;

DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE ESTS D23839(R0339).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0708602.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP001539; BAA92910.1; .
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH_domain; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PS50084; KH_TYPE_1; 4.
 SQ SEQUENCE 774 AA; 84066 MW; 5992CCFE0F74BB56 CRC64;

Query Match 8.9%; Score 262.5; DB 10; Length 774;
 Best Local Similarity 21.8%; Pred. No. 4.8e-09;
 Matches 129; Conservative 73; Mismatches 185; Indels 205; Gaps 20;
 QY 156 DEMAQONPQQPRGRRLGQSGSSROGSPGVSQKQPCDLP-----LRLLYP 203
 Db 57 EDMASTRPSKRP-----FKNSSEQNGRGWKTKHNSLOOQPLIVQGVPIFRILCP 110
 QY 204 TQFVGAILKKGATIRNTKQTSKI-----DVRHKEA 237
 Db 111 TSGSGNVIGKGGIIRKQETGVKIRVDEVPQCDERIVITADKQREVSHVOTKEND 170
 QY 238 GA-----AEKSI----- 244
 Db 171 GGVACSVGNHGMKEDHTKEEDSKENKDDSEKGLKEEDKDDKEDDESGDN 230
 QY 245 -----TILSTPEGTSAACKSILEIMHKEAQDIKFTEE----- 276
 Db 231 DKKADSSVAKDPSNPEPEAQLEKGMPLAVKAILVF-----DRIFVNMENGTGDASG 285
 QY 277 ---IPLTAHNFVRLIGKEGRNLKIEQDTDFKITISPLQELTLYNPRTITVKN 332
 Db 286 ERNHVSLRLVLDSQVGLLGNKNGSVIKQMSDSCCEIRVS-----KDK 329
 QY 333 VETCAKAEIEIMKKIRE---SVENDIASMNLQALH---IPGLNLNALGLF--PPTS----- 380
 Db 330 LPLCALPRDELQOILHQKITIYSLHISFLQITGELDSVRKGLNTVAQLLFTTHPPKESDWL 389
 QY 381 -----GMPP-----PTSGPPSAMTPPYQ-----PQFQSETETVH 409
 Db 390 GAHNSGSSRSFNQDVLPPGMQPNLHLFPQGNVAHLNPFPEALMHGHSVPPPELTFR 449
 QY 410 QFTPALSVGAILKGQGHQIKQLSRFAGASIKIAPAEAPDAKVVRMVIITGPP---EAQFKA 466
 Db 450 LLCSSDKVGGIIGKGGNNIKSIQNDTGCEIKVLDI-VPKSEDRIVIFSGPAHGGIGSPA 508
 QY 467 QGRIGYKIEENFVSPKEVKLEAHTVPSFAGRVIGGKTVNELNLSAEVVPVR-PR 525
 Db 509 QNALHVQRKVIPTNKTREGPAICRLIIVSPNQVGLLGGKGSIIAEMKLSGAHIIIVLSK 568
 QY 526 DQTPD---DENDQVVKITGHFYACQVAKRIQEIITGVQKQHQOQKALOGPP 574
 Db 569 DKIPKGVPEDE-VVQISG---ASEAIEALMQITARLNHLFRDRMASTVP 616

RESULT 15
 Q8S7G1 PRELIMINARY; PRT; 542 AA.
 ID Q8S7G1
 AC Q8S7G1;

```
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative nucleic acid binding protein.
GN OSJNB0048D20.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0048D20 genomic sequence.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084766; AAL82529.1; -
SQ SEQUENCE 542 AA; 57843 MW; B28A54BB2450F336 CRC64;

Query Match 8.8%; Score 260.5; DB 10; Length 542;
Best Local Similarity 22.5%; Pred No. 3.9e-09;
Matches 105; Conservative 94; Mismatches 191; Indels 77; Gaps 17;

QY 168 |PGRRRGLGQR-----GSSRQSGPG-SVSKQKPCDPLRLVPTQFVGAIIGKEGATIR 219
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 |PGRHSHGKRHSYAEANGGKGRNPGDDTVAPGDDTVVRYLCPRSKIGSIIGRGEIAK 63
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 220 NITKQTSKIDVRKENAGAAKSIIYLTPEGTSACKSILEIMHK--EAQDIKF---- 273
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 |QMRADTQAKIRI-GEVSVCGERVITIFSSSRETN-----TLVDAEDKVCPAQDALFRVHE 118
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 274 |-----TEETPLKILAHNNFVRLIGKEGRNLKKIEQDQDITITISPLQ 317
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 119 |KLSIDDDIGNEEDEGLAQVTVRLVPSDQIGCIIGKGHIQGIIRSDTCAHIRVLSNEN 178
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 318 L--TLYNPERTITVKNVETCAKAEIEIMKKIRESYENDIASNNLOA-----HL 364
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 179 |PACATSGDELQITSGDSTVVRKALLQVSSRL---HDNPSRSHLLASSMTQPPYPVGS 235
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 365 |PGLNLNALGLFPPTS--CMPPPTSGPPSAMTPPYQPEQSETEVHQFIPALSVGAI 421
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 236 |GSSSTAPVVGITPLISSYGGYKGDVADWPSIYQPRESSAKESRLRLCRAASNVGGVI 295
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 422 |GKQCQHIKQLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAQGRIVYGIKEENFVS 481
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 296 |GKGGIITKIQIESGAFIKVDSSNTEDD---CIITVSAKEFF--EDPVSPITINAAVHLQ 349
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 482 |KKEVKLEAHIRVPSFAA-----GRVIGKGGTVNELQNLSSAEV-VVPRDQTPD- 530
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 350 |PRGSEKTDPSAIPSYTTRLLVSTSRIGCLIGKGGSIITEIRTSRANIRILSKENVPKV 409
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 531 |ENDOVVKITGHFYACVAQRKIQELTQVKOH--QOQKALQSGPP 574
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 410 |AADEEMVQISGDL---DVVRHALLQITIRKANFEREGALSGFPP 453
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: April 24, 2003, 16:14:14
Job time : 42 secs

GenCore version 5.1.4.p5.4578
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:06 ; Search time 14 Seconds
(without alignments)

1715.342 Million cell updates/sec

Title: US-09-897-778-176

Perfect score: 2956

Sequence: 1 MNKLYIGNLSENAAPSDLES.....VKHQHQKALQSGPQSRRK 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 227 | 7.7 | 510 | 1 NOAL_HUMAN | P51513 homo sapien |
| 2 | 24.5 | 7.6 | 339 | 1 PCB3_HUMAN | P57721 homo sapien |
| 3 | 219.5 | 7.4 | 339 | 1 PCB3_MOUSE | P57722 mus musculus |
| 4 | 213.5 | 7.2 | 403 | 1 PCB4_HUMAN | P57723 homo sapien |
| 5 | 212.5 | 7.2 | 403 | 1 PCB4_MOUSE | P57724 mus musculus |
| 6 | 210.5 | 7.1 | 606 | 1 Y475_ARATH | P58223 arabidopsis |
| 7 | 209.5 | 7.1 | 356 | 1 PCB1_HUMAN | Q15365 homo sapien |
| 8 | 209.5 | 7.1 | 356 | 1 PCB1_RABIT | Q19048 oryctolagus |
| 9 | 207.5 | 7.0 | 1268 | 1 VGLN_HUMAN | Q00341 homo sapien |
| 10 | 204 | 6.9 | 413 | 1 PBP2_YEAST | P38151 saccharomyc |
| 11 | 202.5 | 6.9 | 365 | 1 PCB2_HUMAN | Q15366 homo sapien |
| 12 | 202 | 6.8 | 362 | 1 PCB2_MOUSE | Q61990 mus musculus |
| 13 | 194.5 | 6.6 | 1270 | 1 VGLN_CHICK | P81021 gallus gall |
| 14 | 190 | 6.4 | 463 | 1 ROK_HUMAN | Q07244 homo sapien |
| 15 | 190 | 6.4 | 463 | 1 ROK_RABIT | Q19049 oryctolagus |
| 16 | 190 | 6.4 | 464 | 1 ROK_MOUSE | Q05077 mus musculus |
| 17 | 183 | 6.2 | 1222 | 1 S160_YEAST | Q06105 saccharomyc |
| 18 | 170.5 | 5.8 | 629 | 1 PAB2_ARATH | P42731 arabidopsis |
| 19 | 164.5 | 5.6 | 644 | 1 PAB4_HUMAN | Q13310 homo sapien |
| 20 | 160.5 | 5.4 | 636 | 1 PAB1_MOUSE | P29341 mus musculus |
| 21 | 159.5 | 5.4 | 470 | 1 NR54_HUMAN | Q15233 homo sapien |
| 22 | 159.5 | 5.4 | 636 | 1 PAB1_HUMAN | P11940 homo sapien |
| 23 | 156 | 5.3 | 633 | 1 PABP_YENLA | P20965 xenopus lae |
| 24 | 152.5 | 5.2 | 381 | 1 YBD2_YEAST | P38199 saccharomyc |
| 25 | 150 | 5.1 | 344 | 1 SFR6_HUMAN | Q13247 homo sapien |
| 26 | 147.5 | 5.0 | 747 | 1 FCA_ARATH | O04425 arabidopsis |
| 27 | 145.5 | 4.9 | 414 | 1 NSR1_YEAST | P27476 saccharomyc |
| 28 | 145.5 | 4.9 | 489 | 1 SFR4_MOUSE | Q8ve97 mus musculus |
| 29 | 145 | 4.9 | 414 | 1 NOP3_YEAST | Q01560 saccharomyc |
| 30 | 139.5 | 4.7 | 522 | 1 PAB2_HUMAN | Q15097 homo sapien |
| 31 | 138.5 | 4.7 | 424 | 1 S3B4_HUMAN | Q15427 homo sapien |
| 32 | 135.5 | 4.6 | 359 | 1 ELV2_HUMAN | Q12926 homo sapien |
| 33 | 135.5 | 4.6 | 360 | 1 ELV2_MOUSE | Q60899 mus musculus |

| | | | | | |
|----|-------|-----|------|--------------|---------------------|
| 34 | 135 | 4.6 | 367 | 1 ELV3_HUMAN | Q14576 homo sapien |
| 35 | 135 | 4.6 | 367 | 1 ELV3_MOUSE | Q60900 mus musculus |
| 36 | 135 | 4.6 | 494 | 1 SFR4_HUMAN | Q08170 homo sapien |
| 37 | 134 | 4.5 | 285 | 1 CABA_MOUSE | Q98020 mus musculus |
| 38 | 134 | 4.5 | 307 | 1 SXL_MEGSC | O01671 megascella s |
| 39 | 133 | 4.5 | 373 | 1 ELV4_RAT | O09032 rattus norv |
| 40 | 133 | 4.5 | 380 | 1 ELV4_HUMAN | P26378 homo sapien |
| 41 | 133 | 4.5 | 385 | 1 ELV4_MOUSE | Q61701 mus musculus |
| 42 | 130 | 4.4 | 633 | 1 ROR_HUMAN | Q43390 homo sapien |
| 43 | 129.5 | 4.4 | 6359 | 1 BACC_BACLI | O68008 b bacitraci |
| 44 | 129 | 4.4 | 653 | 1 PABP_SCHPO | P31209 schizosacch |
| 45 | 128.5 | 4.3 | 660 | 1 PAB3_ARATH | O64380 arabidopsis |

ALIGNMENTS

RESULT 1

| | | | | | |
|------------|---|-----------|------|---------|--|
| NOAL_HUMAN | | | | | |
| ID | NOAL_HUMAN | STANDARD; | PRT; | 510 AA. | |
| AC | P51513; | | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | | |
| DE | Onconeural ventral antigen-1 (NOVA-1) (Paraneoplastic Ri antigen) | | | | |
| DE | (Ventral neuron-specific protein 1). | | | | |
| GN | NOVA1. | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo. | | | | |
| OX | NCBI_TaxID=9606; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=Cerebellum, and Hippocampus; | | | | |
| RA | MEDLINE=94000830; Pubmed=8598153; | | | | |
| RA | Buckanovich R.J., Posner J.B., Darnell R.B.; | | | | |
| RT | "Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding | | | | |
| RT | protein and is specifically expressed in the developing motor | | | | |
| RT | system." | | | | |
| RL | Neuron 11:657-672(1993). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE OF 1-34 FROM N.A. | | | | |
| RC | TISSUE=Fetal brain; | | | | |
| RA | Dmitrenko V.V., Garifulin O.M., Shostak K.A., Smikodub A.I., | | | | |
| RA | Kavsan V.M.; | | | | |
| RL | Submitted (APR-1996) to the EMBL/GenBank/DBJ databases. | | | | |
| CC | -1- FUNCTION: MAY REGULATE RNA SPLICING OR METABOLISM IN A SPECIFIC | | | | |
| CC | SUBSET OF DEVELOPING NEURONS. | | | | |
| CC | -1- SUBCELLULAR LOCATION: Nuclear. | | | | |
| CC | -1- TISSUE SPECIFICITY: BRAIN. | | | | |
| CC | -1- DISEASE: ANTIBODIES ARE SEEN AGAINST THIS ANTIGEN IN THE PATIENTS | | | | |
| CC | SUFFERING WITH PARANEOPlastic OPSOCLONUS-ATAXIA (POA). POA IS A | | | | |
| CC | DISORDER IN WHICH ABNORMAL MOTOR CONTROL OF THE EYES, TRUNK AND | | | | |
| CC | LIMBS DEVELOP IN WOMEN WITH BREAST OR SMALL LUNG CANCER. POA | | | | |
| CC | PATIENTS SUFFER FROM OPSOCLONUS, A CHAOTIC EYE MOVEMENT DISORDER | | | | |
| CC | ATTRIBUTABLE TO A LACK OF INHIBITION OF BURST NEURONS IN THE BRAIN | | | | |
| CC | STEM: MYOCLONUS, A DISORDER PERHAPS ATTRIBUTING TO A FAILURE OF | | | | |
| CC | INHIBITORY CONTROL OVER SPINAL MOTOR NEURONS; AND TRUNCAL ATAXIA, | | | | |
| CC | REFERABLE TO CEREBELLAR DYSFUNCTION; THERE ARE NO CORTICAL OR | | | | |
| CC | SENSORY DEFICITS. | | | | |
| CC | -1- SIMILARITY: CONTAINS 3 KH DOMAINS. | | | | |

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U04840; AAA16022.1; --
EMBL: Z70771; CAA94810.1; --
Genew; HGNC:7886; NOVA1.

```

DR MW: 602157;
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type_1.
DR Pfam: PF00013; KH-domain; 3.
DR SMART: SM00322; KH; 3.
DR PROSITE: PS0084; KH_TYPE_1; 3.
DR Nuclear protein; RNA-binding; Repeat; Alternative splicing; Antigen.
KW DOMAIN 52 87
FT DOMAIN 177 212
FT DOMAIN 427 462
FT VARSPLIC 88 90 MISSING (IN TUMOR ISOFORM).
FT VARSPLIC 153 176 MISSING (IN TUMOR ISOFORM).
FT VARSPLIC 177 184 VKIIVNS -> KHLISWIS (IN TUMOR ISOFORM).
FT VARSPLIC 185 510 MISSING (IN TUMOR ISOFORM).
SQ SEQUENCE 510 AA: 52055 MW; 7B9BA867CAB6984 CRC64;

Query Match 7.7%; Score 227; DB 1; Length 510;
Best Local Similarity 22.4%; Pred. No. 5.1e-07;
Matches 119; Conservative 77; Mismatches 179; Indels 156; Gaps 20;

QY 155 PDMAAQNPLQOPRRGRGLGQSSRGSGSKOKPCDPLRLVLPQFVGAITGKE 214
DB 23 PD---SKRPLEAP-----PEAGSKRTNTGEGQYF-----LKVLPVSAAGSIIGK 68

QY 215 GATIRNTTKQTSKIDVHRENA-----GAAEKSTITLSPECTSAAACKSILEIMHKEAQ 269
DB 69 GQIVQLQKETGATIKLSKSKDFYPTTERTVCLIOGTVEALNAVHGFIAEKIREMPQ 128

QY 270 DIKTEEPL-----KILAHNNFVGR 290
DB 129 NVAKTEPVSILQPTQVNDRIKQLPSSPTTKSSSDPMITSRANQVKIIVPNSTAGL 189

QY 291 LIKEGRNKKIIBQDTHKITISPLQELTYNPRTITVKNVETCAKAEIEIMKKIRE- 349
DB 189 IIGRGATVAVMEQSGAWQLSQKPD-GINLQERVVVTVSGEPQNRKAVELIIGKIGED 247

QY 350 -----SYEN---DIASMN-----LQAHILPGLNLNGLGFPPT- 379
DB 248 PQSGCLNISYANVTGVPVANSNPTGSPYANTAEVLPTAAAGLLGHANLACVAAPFAVL 307

QY 380 SGMPPTSGPPSANTPPYQFQSETEVHQFIPALSGAIGKQGHKQLSRFAGASI 439
DB 308 SGF---TGNDLVAIT-----SALNTL-----ASYGNLNTLGLUGLQANA-TGALA 349

QY 440 KIAPAEAPDAKVRWVIITGPPAPQKAGRIYGI----- 474
DB 350 AAAASANPAAAAANLLATYASEA--SASGSTAGGTAGTAFGLSALAAATAATNGYFGAASP 407

QY 475 -----KEENFVSPKEEVKLEAHIRVPSFAAGRVIGKGTVNELOLSSAEVVVPR- 525
DB 408 LAASAILGTETKSDGSKDVE---IAVPENLVGAIGLKGKTLVEYQELTGARTQISK 463

QY 526 -DQTPDENQVVKITGHFYACOVAOKIOEILTQVQHOOKALQSGPPQ 575
DB 464 GEFVPGTRNRKVT-ITGTPAATAQAQYLITQRII-----YEOGVRAANPQ 507

RESULT 2
PCB3_HUMAN
ID PCB3_HUMAN STANDARD; PRT; 339 AA.
AC P57721;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE POLY(rC)-binding protein 3 (Alpha-CP3).
GN PCB3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20356135; PubMed-10936052;

```

```

RA Makeyev A.V., Liehaber S.A.;
RT "Identification of two novel mammalian genes establishes a subfamily
of KH-domain RNA-binding proteins.";
RL Genomics 67:301-316(2000).
CC -1- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT
CC BINDS PREFERENTIALLY TO OLIGO DC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF176329; AAC09240.1; -.
CC HSP: Q07244; IKHM.
CC Genew; HGNC:8651; PCB3.
CC InterPro: IPR004087; KH_dom.
CC InterPro: IPR004088; KH_type_1.
CC Pfam: PF00013; KH-domain; 3.
CC SMART: SM00322; KH; 3.
CC PROSITE: PS0084; KH_TYPE_1; 3.
CC Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
KW Repeat.
KW DOMAIN 13 63 KH 1.
KW DOMAIN 97 150 KH 2.
KW DOMAIN 261 313 KH 3.
FT DOMAIN 13 63
FT DOMAIN 97 150
FT DOMAIN 261 313
SQ SEQUENCE 339 AA; 35938 MW; F9F19FF7590C4188 CRC64;

Query Match 7.6%; Score 224.5; DB 1; Length 339;
Best Local Similarity 25.0%; Pred. No. 4.1e-07;
Matches 83; Conservative 65; Mismatches 119; Indels 65; Gaps 14;

QY 196 LPRLVLPQTFVGAIGREGATIRNITKQTSKIDVHRENAAGAEKSTITLSPECTSA 255
DB 14 LTIRLLHGHKEVSGIIGKGTVMKREESGARINI-----SEGNCPEIRVITGP--TDA 67

QY 256 AKSILEIMHKEAQDI-----KFTTEIP--LKILAHNNFVGRIGLKGNLAKIKEDT 306
DB 68 IFKAFAMIAKFEEDIIISMNSPATSKPPVTLRVLPASQCGSLIGKSGKIREIST 127

QY 307 DTKITISPLQELTYNPRTITVKNVETCAKAEIEIMKKIRES-----YENDIASM 358
DB 128 GAQVQVA--GDMLPNSTERAVTISGTPDAIQCQVKQICVVMLESPPKGTATPYRKPAST 185

QY 359 NL-----QAHILPG-----LNLNALGL-----FPPTSGMPPPTSGPPSAMTPPYQ 399
DB 186 PVIFAGGQAVTIQGOVAIPHDPQLTKLHQLAMQQTFFPLQGTNPAF--PGEKLP LHS 242

QY 400 -----PEQSETEVHQF-IPALSGAIGKQGHKQLSRFAGASIKIAPAEA 446
DB 243 EEAQNLWGSSGLDASPPASTHELTIFNDLIGCIIGRTKINEIFQMSGAQIANTATE 302

QY 447 PDAKVRWVIITGPE-----AQFKAQRIYGI 474
DB 303 GSSE-RQITITGTPTANISLAQYILNARLTSEV 333

RESULT 3
PCB3_MOUSE
ID PCB3_MOUSE STANDARD; PRT; 339 AA.
AC P57722;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE POLY(rC)-binding protein 3 (Alpha-CP3).
GN PCB3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```


| | | | | | | |
|----|-----|-----------------------|---------------|------------------|------------------------|--------------------|
| Db | 18 | LTLRLMHGKEVSLIGKKG | GETVKRI | RESSARITI----- | SEGSGPERITIT'--GSTA | 71 |
| Qy | 256 | ACKSILEIMHKEAADI----- | -KFTIEIP | KLILAHNNEVGR | LICKEGRNLKKIEQDT | 306 |
| Db | 72 | VFHAVSMIAFKLDEDLCA | APANGSVSRPVT | LRVLVPASCGSL | GKAGTKIKEIRETT | 131 |
| Qy | 307 | DTKTIISPLQELTYLNP | ERITVKGWVEYCA | KAEIEMKKIR | ESYENDIASMNLQALIP | 366 |
| Db | 132 | GACQOVA--GDLNP | STERAVTVSGV | PDAILVCVQIC | AVILESPPK---GATIPYH--P | 184 |
| Qy | 367 | GLNLNALGLFPPTSG | MPPTSGPPS | AMTP-----PY----- | -PQEQSETE | 406 |
| Db | 185 | SISLGTV--LLSANQF-- | SVQCGYGA | TPAEVTKLQOLS | GHAVFASPSVPGMDP | STQT 241 |
| Qy | 407 | TVHQF-IPALS | VGAIIGKQGH | IKQLSRFAGAS | IKIAPAEADPAKVR | WVILITGPPEAQFK 465 |
| Db | 242 | SSQEFVLNDLTCV | QSGSKSEIR | QMSGAHIKIG-NO | EGAGERHVTITSP | VSVAL 300 |

Db 301 AQ 302

RESULT 6

Y475_ARATH STANDARD; PRT; 506 AA.

AC P58223; O49507;

AD 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative nucleic acid binding protein At4g18375.

GN AT4G18375 OR F28J12.30.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

ON NCBI

XX SEQUENCE FROM N.A.

RP STRAIN=cv. Columbia;

RC MEDLINE=20083488; Pubmed=10617198;

EX Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N., Harris B., Ansong W., Brandt P., Grivell L., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M., Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T., Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Vandenbussche F., Van der Schueren J., Grynoprez B., Chuang Y.-J., Bastiaens I., Aert R., Defoor E., Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M., Holzer E., Brandt A., Peters S., van Staveren M., Dirke W., Wooljman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P., Bernreiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buysshaert C., Giesen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A., Neumann S., Argirious A., Vitale D., Liguori R., Piravallin E., Massenot O., Quigley F., Clabaud G., Muendlein A., Falber R., Schnabl S., Hillier R., Schmidt W., Lecharny A., Aubourg S., Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E., Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T., Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C., Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Johnson D., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L., Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C., Antonolu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen E., Marra M., Martienssen R., McCombie W.R.; "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."; Nature 402:769-777(1999).

RP CONCEPTUAL TRANSLATION.

RA Kieselbach T.; Unpublished observations (JUL-2001).

RL -!- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -!- SIMILARITY: CONTAINS 5 KH DOMAINS.

CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A GENE PREDICTION ERROR AND THE FUSION OF THE CDS FOR THIS PROTEIN WITH THE CDS FOR CHLOROPLAST PROTEASE HHOA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; AL021710; CAA16717.1; ALT_SEQ.

DR EMBL; AL161548; CAB78839.1; ALT_SEQ.

DR InterPro: IPR004087; KH_dom.

DR InterPro: IPR004088; KH_type_1.

DR Pfam: PF00013; KH-domain; 5.

DR SMART; SM00322; KH; 5.

DR PROSITE; PS50084; KH_TYPE_1; 5.

DR KW Hypothetical protein; Nuclear protein; RNA-binding; Repeat.

FT DOMAIN 35 99 KH 1.

FT DOMAIN 138 210 KH 2.

FT DOMAIN 311 380 KH 3.

FT DOMAIN 394 455 KH 4.

FT DOMAIN 535 599 KH 5.

FT SEQUENCE 606 AA; 65760 MW; 61F135BBB8647C0C CRC64;

Query Match 7.1%; Score 210.5; DB 1; Length 606;

Best Local Similarity 19.6%; Pred. No. 6.8e-06;

Matches 89; Conservative 86; Mismatches 187; Indels 93; Gaps 14;

Qy 199 RLLVPTQFVGAIGKEGATIRNTKQTSKIDVHRKENAGAESTITL----- 247

Db 39 RILCPIDWGVGKGVKGVINAIRHNTKAKIKVPDLHF-GCSQRVITVYGVKQDEIG 97

Qy 248 ---STPEGSACKSILEIM-----HKEAQDIFTEIPKILAHNFVGRLLGKE 295

Db 98 FTKSENEPLCAQDALLKVDVAIVASDEENNTKTVDRDNKCRLLVPFSQSSLLGKA 157

Qy 296 GRNLKKIBODTDTKI-----TISPLQELTYNPERTITVKGNVETCAK---AEIEIMKKI 347

Db 158 GENTKIRIRRRTRASVKVSKVDSPSHVCAMEYDNVVVISGEPSVKQALFAVSAMVKI 217

Qy 348 ---RESYENDIASMNI-----QHL-----IPGLNLAIGL 375

Db 218 NPRENIPLDTSQDVPAAASVIVPSDLNSVYPQTGFYSNQDHIHQAGVPSY-FNALSV 276

Qy 376 -----FPPISGMPPTSGPPSAMTPPYPPQFEQSETEVHQIPALSVGAIGKQGHQIKQ 430

Db 277 SDFQGYAETAANVPVFPASSLPVTHFGGSGSRSEELVFKVLCPLCNMRVICKGSGTIKR 336

Qy 431 LSRFAGASIKIAPAEAPDAKVR-----MWITGPPEAQKQAGRIYGIKEENFVSPKE 484

Db 337 IRASGSCI-----EYVDSRTKCGDDCEVIIVTATESPDDMKSMAVEAVLLQLQYINDED 391

Qy 485 EVKLEAHIRVPSFAAGVIGKGTQVNEONLSSAEVVPVPROTTPDENQVVKITGHFY 544

Db 392 AENVKMLLVSSKVGICVSGVINEIRKTNANICISKGRKDD-----LVEVSGEV- 445

Qy 545 ACQVARKIKQIEILTVQKHOQKALQSGPPQSRKK 579

Db 446 -----SSVRDALIQIVLRREDVGLGKDSVATRK 474

RESULT 7

PCBL_HUMAN

ID PCBL_HUMAN STANDARD; PRT; 356 AA.

AC Q15365; Q13157; Q14975;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Poly(rC)-binding protein 1 (Alpha-Cp1) (hnrNP-E1) (Nucleic acid binding protein SUB2.3).

DE PCBP1.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95331278; PubMed=7607214;

RA Leifers H., Dejgaard K., Celis J.E.;

RT "Characterisation of two major cellular poly(rC)-binding human proteins, each containing three K-homologous (KH) domains.";

RL Eur. J. Biochem. 230:447-453(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=96016208; PubMed=7556077;

RA Killedjian M., Wang X., Liebhauer S.A.;

RT "Identification of two KH domain proteins in the alpha-globin mRNA stability complex.";

RL EMBO J. 14:4357-4364(1995).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Lymphocytes;

RX MEDLINE=94203810; PubMed=8152927;

RA Aasheim H.-C., Loukianova T., Deggerdal A., Smeland E.B.;

RT "Tissue specific expression and cDNA structure of a human transcript encoding a nucleic acid binding [oligo(dc)] protein related to the pre-mRNA binding protein K.";

RL Nucleic Acids Res. 22:959-964(1994).

CC -!- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT BINDS PREFERENTIALLY TO OLIGO DC.

CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND IN THE NUCLEUS. MAY SHUTTLE BETWEEN THE NUCLEUS AND THE CYTOPLASM.

CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN SKELETAL MUSCLE, THYMUS AND PERIPHERAL BLOOD LEUCOCYTES WHILE A LOWER EXPRESSION IS OBSERVED IN PROSTATE, SPLEEN, TESTIS, OVARY, SMALL INTESTINE, HEART, LIVER, ADRENAL AND THYROID GLANDS.

CC -!- PTM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED THE STRONGEST POLY(RC)-BINDING ACTIVITY.

CC -!- SIMILARITY: CONTAINS 3 KH DOMAINS.

CC -!- CAUTION: THERE IS PROBABLY A FRAMESHIFT ERROR IN THE NUCLEOTIDE SEQUENCE OF REF.3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; X78137; CAA55016.1; -

DR EMBL; U24223; AAA91317.1; -

DR EMBL; Z29505; CAA82631.1; ALT_FRAME.

```
DR HSP; Q07244; IKHM.
DR Genew; HNC:8647; PCBP1.
DR MIN; 601209;
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH-domain; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS00084; KH_TYPE_1; 3.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
KW Phosphorylation; Repeat.
FT DOMAIN 13 75 KH 1.
FT DOMAIN 97 162 KH 2.
FT DOMAIN 279 343 KH 3.
FT CONFLICT 205 205 V -> A (IN REF. 2 AND 3).
FT SEQUENCE 356 AA; 37526 MW; DC85477576D5104 CRC64;
SQ
Query Match 7.1%; Score 209.5; DB 1; Length 356;
Best Local Similarity 22.2%; Pred. No. 3.8e-06;
Matches 80; Conservative 69; Mismatches 110; Indels 101; Gaps 13;
QY 196 LPLRLVPTQFVGAIGKEGATIRNTKQSKIDVHRKENAGAAEKSTITLSTPEGTS 255
DB 14 LTIRLLMHGKEVGSIIKGKESVKRIRESGARINI-----SEGNCPERIITLTGP--TNA 67
QY 256 ACKSILEIMHKEAQDIKFT-----EIPLKILAHNNFVRLIGKEGNLKKIEQDT 306
DB 68 IFKAFAMIDKLEEDINSSMTNSTAASRPVTLRLVWPATQCGSLIGKGGCKIKIREST 127
QY 307 DTKITISPLQELTYLPERTITVKNVETCAKAEIEI---MKKIRESYENDIASMNLQA 362
DB 128 GAQVQVA--GDMLPNSTERAITIAGVPQSVTECVKQICLVMLETUSQSGRVRMTIPYQP 185
QY 363 HLIPGLNLALGLFPPTS-----GMPPPT---SGPP----- 390
DB 186 -----MPASSPVICAGGQDRCSDAAGYPHATHDLEGPLDAYSIOGHTISP 232
QY 391 -----SMTPTPYQFE-----QSETETVHQF--IPALSVG 418
DB 233 LDIAKLNOVARQOSHFMHGGTGAGIDSSPEVKGWASLDASTQTTHLTIPNNLIG 292
QY 419 AIIGKQGHQTKQISRFAGASIKTA--PAEAPDAKVRMVIITGPPEAKFKAQRIYGIKEE 477
DB 293 CIIGROGANINEIROMSGAQIKIANPVEGSG--RQVTITGSAASISLAQYLINARLSSE 350
RESULT 8
PCBL_RABIT
AC O19048;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly(rC)-binding protein 1 (Alpha-CPI) (hnrNP-E1).
GN PCBP1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99207122; PubMed=10101190;
RA Thiele B.J., Berger M., Huth A., Reimann I., Schwarz K., Thiele H.;
RT "Tissue-specific translational regulation of alternative rabbit
RL 15-lipoxygenase mRNAs differing in their 3'-untranslated regions.";
RL Nucleic Acids Res. 27:1828-1836(1999).
CC CC -1- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT BINDS
CC CC REFERENTIALLY TO OLIGO DC (BY SIMILARITY).
CC CC -1- SUBCELLULAR LOCATION: Nuclear (BY SIMILARITY).
CC CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ003023; CAA05814.1;
CC HSP; Q07244; IKHM.
CC InterPro; IPR004087; KH_dom.
CC InterPro; IPR004088; KH_type_1.
CC Pfam; PF00013; KH-domain; 3.
CC SMART; SM00322; KH; 3.
CC PROSITE; PS00084; KH_TYPE_1; 3.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding; Repeat.
KW Nuclear protein; RNA-binding; KH 1.
FT DOMAIN 13 75 KH 1.
FT DOMAIN 97 162 KH 2.
FT DOMAIN 279 343 KH 3.
FT SEQUENCE 356 AA; 37497 MW; 6D1A261276CA206D CRC64;
SQ
Query Match 7.1%; Score 209.5; DB 1; Length 356;
Best Local Similarity 22.8%; Pred. No. 3.8e-06;
Matches 79; Conservative 70; Mismatches 123; Indels 75; Gaps 12;
QY 196 LPLRLVPTQFVGAIGKEGATIRNTKQSKIDVHRKENAGAAEKSTITLSTPEGTS 255
DB 14 LTIRLLMHGKEVGSIIKGKESVKRIRESGARINI-----SEGNCPERIITLTGP--TNA 67
QY 256 ACKSILEIMHKEAQDIKFT-----EIPLKILAHNNFVRLIGKEGNLKKIEQDT 306
DB 68 IFKAFAMIDKLEEDINSSMTNSTAASRPVTLRLVWPATQCGSLIGKGGCKIKIREST 127
QY 307 DTKITISPLQELTYLPERTITVKNVETCAKAEIEI---MKKIRESYENDIASMNLQA 361
DB 128 GAQVQVA--GDMLPNSTERAITIAGVPQSVTECVKQICLVMLETUSQSGRVRMTIPYQP 185
QY 362 AHLIPGLNLALGLFPPTS-----SGPP----- 390
DB 186 MPASSPVICAGGQDRCSDAAGYPHATHDLEGPLDAYSIOGHTISPDLAKLNOVARQ 245
QY 391 -----SMTPTPYQFE-----QSETETVHQF--IPALSVGAIGKQGHQTKQL 431
DB 246 SHFAMHGGTGAGIDSSPEVKGWASLDASTQTTHLTIPNNLIGCIIRQGANINEI 305
QY 432 SRFAGASIKTA--PAEAPDAKVRMVIITGPPEAKFKAQRIYGIKEE 477
DB 306 RQMSGAQIKIANPVEGSG--RQVTITGSAASISLAQYLINARLSSE 350
RESULT 9
VGLN_HUMAN
ID VGLN_HUMAN STANDARD; PRT; 1268 AA.
AC Q00341; Q9UCY3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vigilin (high density lipoprotein-binding protein) (HDL-binding
DE protein).
DE
GN HDLBP OR HBP OR VGL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92291094; PubMed=1318310;
RA McKnight G.L., Reasoner J., Gilbert T., Sundquist K.O., Hokland B.,
RA McKernan P.A., Champagne J., Johnson C.J., Bailey M.C., Holly R.,
RA O'Hara P.J., Oram J.F.;
RT "Cloning and expression of a cellular high density lipoprotein-binding
RT protein that is up-regulated by cholesterol loading of cells.";
RL J. Biol. Chem. 267:12131-12141(1992).
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=96184515; PubMed=8605996;
```



```

Db 129 MIVRALEEHNEDNGEDI-----EISINLLIPIHLMGCIIGRGRSLRIEDLSRAKFLPA 184
QY 313 SPLQELTYLNPERTITVKGNETVCARAKAEIEMKIRIESYENDIASNMNQAHILPGLNLNA 372
Db 185 SPNQ--LLLSNDRILINGVPDAIHATFYISQTL-----LNFQME-SPQKNVKR 231
QY 373 LGLFPPTSGMPPPTSGPPSAMTPPYQPF-----EQSETETV-----HQFIPALSVGAIIIG 422
Db 232 SIYYQPT-----QFNSVLIDHSQNTIFHQRNHQYHP--SDKLLSY 270
QY 423 KOQHILKQLS-----RFAGASIKIAPAEAPDAKV-RMWIITGPEAQFRAQGRIVYK 473
Db 271 KPNKNLPISSTLLSMATPOYTASVANATAFQPNFVNPVTVLDGVPVISPAPGNHLLMNF 330
QY 474 IKEENFVSPEKEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVVPRDQTPDE-N 532
Db 331 VQOEIIFDEK-----FVGNVIGDKGHINSKVESTGCSIII---QDPVEGS 373
QY 533 DQVVVKITGHFYACOYA 549
Db 374 SERRLTIRGTFFMASQAA 390

RESULT 11
PCB2_HUMAN STANDARD; PRT; 365 AA;
AC Q15366;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly(rC)-binding protein 2 (Alpha-CP2) (hnrnp-E2).
GN PCBP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95331278; PubMed=7607214;
RA Jeffers H., DeJgaard K., Cellis J.E.;
RT "Characterisation of two major cellular poly(rC)-binding human
RT proteins, each containing three K-homologous (KH) domains.";
RL Eur. J. Biochem. 230:447-453(1995).
CC -1- FUNCTION: MAJOR CELLULAR POLY(RC)-BINDING PROTEIN. BINDS ALSO
CC POLY(RU).
CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND IN THE NUCLEUS. MAY SHUTTLE
CC BETWEEN THE NUCLEUS AND THE CYTOPLASM.
CC -1- TISSUE SPECIFICITY: DETECTED IN ALL TISSUES EXAMINED.
CC -1- PTM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED THE
CC STRONGEST POLY(RC)-BINDING ACTIVITY.
CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X78136; CAAS5015.1; --
CC DR HSSP; Q07244; 1KHM.
CC DR Genew; HGNC:8648; PCBP2.
CC DR MIM; 601210; --
CC DR InterPro; IPR004087; KH_dom.
CC DR InterPro; IPR004088; KH_type_1.
CC DR Pfam; PF00013; KH-domain; 3.
CC DR SMART; SM00322; KH; 3.
CC DR PROSITE; PS00084; KH_TYPE_1; 3.
CC DR Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
CC phosphorylation; Repeat.
CC DOMAIN 13 75 KH 1.
FT
```

```

ET DOMAIN 97 162 KH 2.
ET DOMAIN 287 351 KH 3.
SO SEQUENCE 365 AA; 39580 MW; 43F035D76FDC2C63 CRC64;

Query Match 6.9%; Score 202.5; DB 1; Length 365;
Best Local Similarity 24.0%; Pred. No. 1.1e-05;
Matches 87; Conservative 58; Mismatches 119; Indels 99; Gaps 13;

QY 196 LPLRLVPTQFVGALIGKEGATIRNIYQTSKIDVHRKENAGAAEKSIITLSTPECTSA 255
Db 14 LTIILLMHGKEVSGIIGKGSVKMREESGARINI-----SEGNCPERIITLAGP--TNA 67
QY 256 ACKSTLEIMHREAOIDIKPT-----EEIPLKILAHNNFVRLIGKEGNLKKIKPDQT 305
Db 68 LKAFAMIIDKLEEDISSMTNSTAASRPVTLRLVVPASOCGSLICKGCKIKEIREST 127
QY 307 DTKITISPLQBELTYLNPERTITVKG---NVTCAKAEIEIMKKI----- 347
Db 128 GAQVQVA--GDMLPNSTERAITIAGIPQSIIECVKQICVVMLETLSQSPKPGVTIPYRPK 185
QY 348 -----RESYENDIASMNLQAHILPGLNLNALGLFPPTSGMP----- 383
Db 186 PSSSPVIFAGGODRYSTGDSASF---PHTTPSMCLN-----PDLEGPPEAYTIQOYA 237
QY 384 -----PPTSG-----PFSAMTPPYQFQESQETETVHQF--IPAL 415
Db 238 IPQPDLTKLHLQAMQOSHFPMTHTGTGTFSGTSSPEVKGYWGLDASAQTTSHELTPND 297
QY 416 SVGAIIGKQGHILKQLSRFAGASIKIA-PAEAPDAKVRMWIITGPPPAQKAGRIYKGI 474
Db 298 LIGCIIGKQGHAKINEIROMSGAQIKIANPVEG---STDQVITITGSAASISLAQYLINVL 355
QY 475 KEE 477
Db 356 SSE 358

RESULT 12
PCB2_MOUSE STANDARD; PRT; 362 AA.
AC Q61990; Q61383; Q62042;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Poly(rC)-binding protein 2 (Alpha-CP2) (Putative heterogeneous nuclear
DE ribonucleoprotein X) (hnrnp X) (CTBP) (CBP).
GN PCBP2 OR HNRNPX OR HNRPX OR CBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=93376518; PubMed=8367306;
RA Hamm K.B., Kim G., Turch C., Smale S.T.;
RT "Isolation of a murine gene encoding a nucleic acid-binding protein
RT with homology to hnrnp X.";
RL Nucleic Acids Res. 21:3894-3894(1993).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (ISOFORM 2).
RX STRAIN=C57BL/6 X 129/Ola; TISSUE=Liver;
RX MEDLINE=94268912; PubMed=8208614;
RA Goller M., Funke B., Gehe-Becker C., Kroeger B., Lottspeich F.,
RA Horak I.;
RT "Murine protein which binds preferentially to oligo-C-rich single-
RT stranded nucleic acids.";
RL Nucleic Acids Res. 22:1885-1889(1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RA Horak I.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
```

SEQUENCE FROM N.A. (ISOFORM 1).
 RA Makeyev A.V., Liehaber S.A.;
 RT Identification of two novel mammalian genes establishes a subfamily
 of KH-domain RNA-binding proteins.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT
 BINDS PREFERENTIALLY TO OLIGO DC.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: CONTAINS 3 KH DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; L19661; AAA03705.1; -;
 DR EMBL; X75947; CAA53546.1; -;
 DR EMBL; X797982; CAA66619.1; -;
 DR EMBL; AF236845; AAK14059.1; -;
 DR EMBL; AF236842; AAK14059.1; JOINED.
 DR EMBL; AF236843; AAK14059.1; JOINED.
 DR EMBL; AF236844; AAK14059.1; JOINED.
 DR HSSP; Q07244; 1KHW.
 DR MGO; MGI:108202; pcbbp2.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH-domain; 3.
 DR SMART; SM00322; KH; 3.
 DR PROSITE; PSS0084; KH_TYPE_1; 3.
 KW Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
 KW Repeat; Alternative splicing. KH 1.
 FT DOMAIN 13 175
 FT DOMAIN 97 162
 FT DOMAIN 284 348
 FT VARSPLIC 194 224 MISSING (IN ISOFORM 2).
 FT VARSPLIC 263 275 MISSING (IN ISOFORM 3).
 SQ SEQUENCE 362 AA; 38221 MW; 70C9AF710E3BF3C0 CRC64;
 Query Match 6.8%; Score 202; DB 1; Length 362;
 Best Local Similarity 24.2%; Pred. No. 1.1e-05;
 Matches 87; Conservative 60; Mismatches 117; Indels 96; Gaps 14;
 QY 196 LPRLVPTQFVGAIIGKCATIRNITKQTSKIDVHRKENAGAEKSTITILSTPEGTA 255
 DB 14 LTRLLMHGREGVSIIGKKGESVKMREESGARINI-----SEGCNPERIITLAGP--TNA 67
 QY 256 ACKSILEIMHKEADIKET-----EEIPLKILAHNNFVRLIGKEGNLKKIEODT 306
 DB 68 IFKAFAMIDKLEEDISSWNTNSTAASRPVTLRLVVPASOGSLIGKGGCKIKIREST 127
 QY 307 DTKITISPLEQLYLPERTIVKG--NVETCAKAEIEIMKI----- 347
 DB 128 GAOVOVA--GDMFLPNSTERAITAGIPOSIECKVQICVVMLESPPKGVITPRPKPSS 185
 QY 348 -----RESYENDIASNNQAHLPGLNLNALGLFPPTSGMP----- 383
 DB 186 PVIFAGGODRYTSGDSASF---PHITPSMCLN-----PDLEGPLEAYITQGOYAIPOQ 237
 QY 384 -----PPTSGPP--SAMTPPVQ-----FEQSETVTHQF--IPALSVG 418
 DB 238 DITKLHQAQQSHFPMTHGNFTGFSGIESSPKGVYAGLDASAQTTSHLITPNDLIG 297
 QY 419 AIIKQGOHILKOLSFAGASIKIA--PAEAPDAKVRMVIITGPPEAOFKAQGRIVGKKEE 477
 DB 298 CIIROGAKINEIROMSGAQIKIANPVEG--STDQVTTTSGAASISIAOYLINVLRSSE 355
 RESULT 13

VGLN_CHICK STANDARD; PRT; 1270 AA.
 ID VGLN_CHICK
 AC P81021;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Vigilin.
 GN HDLBP OR VGL.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC NCBI_TaxID=9031;
 RN [1]
 CC SEQUENCE FROM N.A.
 CC TISSUE=Embryonic sternal cartilage;
 CC MEDLINE=9229898; PubMed=1606952;
 RA Schmidt C., Henkel B., Poeschl E., Zorbas H., Purschke W.G.,
 RA Gloor T.R., Mueller P.K.;
 RA "Complete cDNA sequence of chicken vigilin, a novel protein with
 amplified and evolutionary conserved domains.";
 RL Eur. J. Biochem. 206:625-634(1992).
 CC -!- SUBCELLULAR LOCATION: INTRACELLULAR.
 CC -!- SIMILARITY: CONTAINS 14 KH DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; X65292; CAA46387.1; -;
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH-domain; 14.
 DR SMART; SM00322; KH; 14.
 DR PROSITE; PSS0084; KH_TYPE_1; 14.
 KW RNA-binding; Repeat.
 FT DOMAIN 150 188
 FT DOMAIN 219 260 KH 1.
 FT DOMAIN 291 333 KH 2.
 FT DOMAIN 360 402 KH 3.
 FT DOMAIN 431 473 KH 4.
 FT DOMAIN 504 545 KH 5.
 FT DOMAIN 577 619 KH 6.
 FT DOMAIN 651 693 KH 7.
 FT DOMAIN 724 766 KH 8.
 FT DOMAIN 798 840 KH 9.
 FT DOMAIN 872 913 KH 10.
 FT DOMAIN 970 1012 KH 11.
 FT DOMAIN 1051 1093 KH 12.
 FT DOMAIN 1126 1168 KH 13.
 FT DOMAIN 1270 AA; 142220 MW; 0A8844F91F9B8619 CRC64;
 SQ SEQUENCE 1270 AA; 142220 MW; 0A8844F91F9B8619 CRC64;
 Query Match 6.8%; Score 194.5; DB 1; Length 1270;
 Best Local Similarity 21.8%; Pred. No. 0.00019;
 Matches 97; Conservative 85; Mismatches 168; Indels 95; Gaps 17;
 QY 182 QGSPGSV--SKQPCDPLR-----LLVPTQFVGAIIGKCATIRNITKQTSKI-- 229
 DB 483 EGOPOGVQQAQKLELLEASRNENERTKDIIIEQKFRHTIIIGKGERIREKFPFEVIIN 542
 QY 230 ---DVHRKENAGAEKSTITILSTPEGTAACKSILEIMHKEADIKFTETIPLKILAHNN 286
 DB 543 FPOPAKSD-----IVQLRGFKNEVEKCTYKQKMWADLVENSFSISVPIKQPHN 594
 QY 287 FVGRLLGKEGNLKKIEODTDTKITISPLEQLYLPERTIVKGNVETCAKAEIEIMKK 346
 DB 595 -----IIGKGANIKIRESNKIDLPGEQAT-----ORQLLSQGRQIVKLLRHRL-- 644
 QY 347 IRESYENDIASNNQAHLPGLNLNAL-----GLTF-----PTSG----- 381

Db 645 ---AOKELANITEVEVSIFSKLHNSLIGAKGRFIRSMEECGVHHFFTEGSGSATVT 701
 Qy 382 -MPPPTSGPPSAMTPYFOFQESQETV-----HQFIPALSYGALIGQGOHIKQL 431
 Db 702 IRAQPRTRWRKPRSSCTWABEKOTKSTYVDLRAKPEVHKF-----LIGRGGGNIRKV 753
 Qy 432 SRPAQKTIAPAPAKVAVVILTPPEAQFRAQGRYIKYKE-ENFVSPKEEVKLEA 490
 Db 754 RONTGARIIFTSEKDOE--LITINGTEEAQVKEAQLKALKNLDNVVEDSMVVDPKH 811
 Qy 491 HIRVFSRAGRVRIGKGTWELQNLSSAEVVVPDTPDENVVVKITGHEYACOVAQ 550
 Db 812 H-----RHVIRRQVLREIADREYGVVRLPTVSGTOSDKVTLK--GAKDCVEAK 861
 Qy 551 RKIQELITVQKHQKQALQSGPPQ 575
 Db 862 KRIQELIEDL---EAQVTICTIPQ 883

RESULT 14
 ROK_HUMAN
 ID Q07244; Q15671; Q96762; PRT: 463 AA.
 AC 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DE Heterogeneous nuclear ribonucleoprotein K (hnRNP K) (DC-stretch
 binding protein) (CSBP) (Transformation upregulated nuclear protein)
 DE (TUNP).
 GN HNRPK OR HNRNPK.
 OS Homo sapiens (Human), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606, 10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA MEDLINE=92107165; PubMed=1729596;
 RA Matunis M.J., Michael W.M., Dreyfuss G.;
 RT "Characterization and primary structure of the poly(C)-binding
 RT heterogeneous nuclear ribonucleoprotein complex K protein.";
 RL Mol. Cell. Biol. 12:164-171(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA MEDLINE=94149726; PubMed=8107114;
 RA DeGaard K., Jeffers H., Rasmussen H.H., Madsen P., Kruse T.A.,
 RA Gesser B., Nielsen H., Celis J.E.;
 RT "Identification, molecular cloning, expression and chromosome mapping
 RT of a family of transformation upregulated hnRNP-K proteins derived by
 RT alternative splicing.";
 RL J. Mol. Biol. 236:33-48(1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC SPECIES=Human, TISSUE=Muscle,
 RA Strausberg K.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; STRAIN=Donryu; TISSUE=Liver;
 RA MEDLINE=94173662; PubMed=8127654;
 RA Ito K., Sato K., Endo H.;
 RT "Cloning and characterization of a single-stranded DNA binding
 RT protein that specifically recognizes deoxycytidine stretch.";
 RL Nucleic Acids Res. 22:53-58(1994).
 RN [5]
 RP MASS SPECTROMETRY.
 RC SPECIES=Human, TISSUE=Breast cancer;
 RA MEDLINE=21829512; PubMed=11840567;
 RA Harris R.A., Tang A., Stein R.C., Lucy K., Brusten L., Herath A.,
 RA Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,

RA zvelebil M.J.;
 RT "Cluster analysis of an extensive human breast cancer cell line
 RT protein expression map database.";
 RL Proteomics 2:212-223(2002).
 RN [6]
 RP STRUCTURE BY NMR OF 375-463.
 RC SPECIES=Human;
 RA MEDLINE=99299390; PubMed=10369774;
 RA Baber J.L., Libutti D., Levens D., Tjandra N.;
 RT "High precision solution structure of the C-terminal KH domain of
 RT heterogeneous nuclear ribonucleoprotein K, a c-myc transcription
 RT factor.";
 RL J. Mol. Biol. 289:949-962(1999).
 CC -!- FUNCTION: ONE OF THE MAJOR PRE-MRNA-BINDING PROTEINS. BINDS
 CC TENACIOUSLY TO POLY(C) SEQUENCES. LIKELY TO PLAY A ROLE IN
 CC THE NUCLEAR METABOLISM OF HNRNAS, PARTICULARLY FOR PRE-MRNAS THAT
 CC CONTAIN CYTIDINE-RICH SEQUENCES. CAN ALSO BIND POLY(C) SINGLE-
 CC STRANDED DNA.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR; NUCLEOPLASM.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS, 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- PTM: PHOSPHORYLATED.
 CC -!- MASS SPECTROMETRY: MW=50976.25; METHOD=MALDI.
 CC -!- SIMILARITY: CONTAINS 3 KH DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S74678; AAB20770.1;
 DR EMBL: X72727; CAA51267.1;
 DR EMBL: BC000355; AAH00355.1;
 DR EMBL: BC014980; AAH14980.1;
 DR EMBL: D17711; BAA04566.1;
 DR PIR: A42058; A42058;
 DR PIR: S41495; S41495;
 DR PIR: IKHM.12-JAN-00.
 DR TRANSFAC: T02853;
 DR SWISS-2DPAGE: Q07244; HUMAN.
 DR Genew: HGNC:5044; HNRPK.
 DR MIM: 600712;
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam: PF00013; KH-domain; 3.
 DR SMART: SM00322; KH; 3.
 DR PROSITE: PS50084; KH_TYPE_1; 3.
 DR Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat; DNA-binding;
 KW Phosphorylation; Alternative splicing; 3D-structure.
 FT DOMAIN 42 92
 FT KH 1.
 FT DOMAIN 144 197
 FT KH 2.
 FT DOMAIN 236 273
 FT RNA-BINDING (RGG-BOX).
 FT DOMAIN 387 439
 FT KH 3.
 FT DOMAIN 54 421
 FT 2 X 22 AA APPROXIMATE REPEATS.
 FT REPEAT 54 76
 FT 1-1.
 FT REPEAT 399 421
 FT 1-2.
 FT DOMAIN 245 329
 FT 2 X 6 AA APPROXIMATE REPEATS.
 FT REPEAT 245 250
 FT 2-1.
 FT REPEAT 324 329
 FT 2-2.
 FT DOMAIN 59 407
 FT 5 X 4 AA REPEATS OF G-X-G-G.
 FT REPEAT 59 62
 FT 3-1.
 FT REPEAT 257 260
 FT 3-2.
 FT REPEAT 267 270
 FT 3-3.
 FT REPEAT 295 298
 FT 3-4.
 FT REPEAT 404 407
 FT 3-5.
 FT DOMAIN 289 294
 FT POLY-PRO.
 FT DOMAIN 310 315
 FT POLY-PRO.
 FT VARSPIC 459 463
 FT SGKFF -> ADVEGF (IN ISOFORM 2).
 FT CONFLICT 32 32 A -> D (IN RSF 2).
 FT SEQUENCE 463 AA; 50976 MW; 0F70BE169B2A064A CRC64;
 SQ


```
Query Match      6.48; Score 190; DB 1; Length 463;
Best Local Similarity 20.78; Pred. No. 8.9e-05;
Matches 87; Conservative 54; Mismatches 131; Indels 148; Gaps 13;

QY 198 LRLVPTQVGAIGKEGATIRNITKQTSKIDVHRKENAGAAEKSTIILSTPEGTSAA 257
DB 45 LRLILQSNAGAVIGKGNKALRTDYNASVS---PDSSGPERILISADIETIGEIL 101

QY 258 KSILEIMHKEAQ--DIKTEEIP-----LKLAHNNFVGRIG 293
DB 102 KKIITPLEEGLQSPATISQPLESDAVECLNYOHFKGSDFCLELRLLIHQSLAGGIIG 161

QY 294 KEGRLKKIEQDTDKTIISPLQELTYNPETITVKGNETCAKAEIEIMKKIRES--- 350
DB 162 VKGAKIKELRENTQT--TIKLFQECCHSTDRVVLIGKPDVVVECIKILDLISESPIK 219

QY 351 -----YENDIASMNLQAHILPGLNINLALGLF---PPTSG---MPP----- 384
DB 220 GRAQPYDPNFYDGYGFTMMFDDRRGPRVGFPMRGRGGFDRMPGGRGPRMPPSRRD 279

QY 385 -----PTSGPPS-----AMTPPPYPO-----F 400
DB 280 YDMSPRRGPPPPPPGGRGGRGSRARNLPLPPPPPRGDLMAYDRRGPRGDRYDGMVGF 339

QY 401 EQSET-----ETVHQFIPALSV 417
DB 340 SADETWDSDAIDTWSPEWQMAVEPOGGSGYDYAGRGSGYCDLGGPIITQTQVTPKDLA 399

QY 418 GAIIGKQGHIIKOLSRFAGASIKI-APAEAPDAKVRMVIITPPPAQFKAQRIYGIKKE 476
DB 400 GSIIGKGGORIKQIRHESGASIKIDEPLG--SEDRIIITGTQDQIQNAQYLLQNSVKQ 457

Query Match      6.48; Score 190; DB 1; Length 463;
Best Local Similarity 20.78; Pred. No. 8.9e-05;
Matches 87; Conservative 54; Mismatches 131; Indels 148; Gaps 13;

QY 198 LRLVPTQVGAIGKEGATIRNITKQTSKIDVHRKENAGAAEKSTIILSTPEGTSAA 257
DB 45 LRLILQSNAGAVIGKGNKALRTDYNASVS---PDSSGPERILISADIETIGEIL 101

QY 258 KSILEIMHKEAQ--DIKTEEIP-----LKLAHNNFVGRIG 293
DB 102 KKIITPLEEGLQSPATISQPLESDAVECLNYOHFKGSDFCLELRLLIHQSLAGGIIG 161

QY 294 KEGRLKKIEQDTDKTIISPLQELTYNPETITVKGNETCAKAEIEIMKKIRES--- 350
DB 162 VKGAKIKELRENTQT--TIKLFQECCHSTDRVVLIGKPDVVVECIKILDLISESPIK 219

QY 351 -----YENDIASMNLQAHILPGLNINLALGLF---PPTSG---MPP----- 384
DB 220 GRAQPYDPNFYDGYGFTMMFDDRRGPRVGFPMRGRGGFDRMPGGRGPRMPPSRRD 279

QY 385 -----PTSGPPS-----AMTPPPYPO-----F 400
DB 280 YDMSPRRGPPPPPPGGRGGRGSRARNLPLPPPPPRGDLMAYDRRGPRGDRYDGMVGF 339

QY 401 EQSET-----ETVHQFIPALSV 417
DB 340 SADETWDSDAIDTWSPEWQMAVEPOGGSGYDYAGRGSGYCDLGGPIITQTQVTPKDLA 399

QY 418 GAIIGKQGHIIKOLSRFAGASIKI-APAEAPDAKVRMVIITPPPAQFKAQRIYGIKKE 476
DB 400 GSIIGKGGORIKQIRHESGASIKIDEPLG--SEDRIIITGTQDQIQNAQYLLQNSVKQ 457
```

```
DR HSSP; Q07244; LKHM.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH-domain; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS00084; KH_TYPE_1; 3.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat; DNA-binding;
KW Phosphorylation.
FT DOMAIN 42 92 KH 1.
FT DOMAIN 144 197 KH 2.
FT DOMAIN 236 273 RNA-BINDING (RGX-BOX).
FT DOMAIN 387 439 KH 3.
FT DOMAIN 54 421 2 X 22 AA APPROXIMATE REPEATS.
FT REPEAT 54 76 1-1.
FT REPEAT 399 421 1-2.
FT DOMAIN 245 329 2 X 6 AA REPEATS OF D-R-R-G-R-P.
FT REPEAT 245 250 2-1.
FT REPEAT 324 329 2-2.
FT DOMAIN 59 407 5 X 4 AA REPEATS OF G-X-G-G.
FT REPEAT 257 260 3-1.
FT REPEAT 267 270 3-2.
FT REPEAT 295 298 3-3.
FT REPEAT 404 407 3-4.
FT DOMAIN 289 294 3-5.
FT DOMAIN 308 315 POLY-PRO.
SQ SEQUENCE 463 AA; 50960 MW; A62A92E645EFE1B6 CRC64;

Query Match      6.48; Score 190; DB 1; Length 463;
Best Local Similarity 20.78; Pred. No. 8.9e-05;
Matches 87; Conservative 54; Mismatches 131; Indels 148; Gaps 13;

QY 198 LRLVPTQVGAIGKEGATIRNITKQTSKIDVHRKENAGAAEKSTIILSTPEGTSAA 257
DB 45 LRLILQSNAGAVIGKGNKALRTDYNASVS---PDSSGPERILISADIETIGEIL 101

QY 258 KSILEIMHKEAQ--DIKTEEIP-----LKLAHNNFVGRIG 293
DB 102 KKIITPLEEGLQSPATISQPLESDAVECLNYOHFKGSDFCLELRLLIHQSLAGGIIG 161

QY 294 KEGRLKKIEQDTDKTIISPLQELTYNPETITVKGNETCAKAEIEIMKKIRES--- 350
DB 162 VKGAKIKELRENTQT--TIKLFQECCHSTDRVVLIGKPDVVVECIKILDLISESPIK 219

QY 351 -----YENDIASMNLQAHILPGLNINLALGLF---PPTSG---MPP----- 384
DB 220 GRAQPYDPNFYDGYGFTMMFDDRRGPRVGFPMRGRGGFDRMPGGRGPRMPPSRRD 279

QY 385 -----PTSGPPS-----AMTPPPYPO-----F 400
DB 280 YDMSPRRGPPPPPPGGRGGRGSRARNLPLPPPPPRGDLMAYDRRGPRGDRYDGMVGF 339

QY 401 EQSET-----ETVHQFIPALSV 417
DB 340 SADETWDSDAIDTWSPEWQMAVEPOGGSGYDYAGRGSGYCDLGGPIITQTQVTPKDLA 399

QY 418 GAIIGKQGHIIKOLSRFAGASIKI-APAEAPDAKVRMVIITPPPAQFKAQRIYGIKKE 476
DB 400 GSIIGKGGORIKQIRHESGASIKIDEPLG--SEDRIIITGTQDQIQNAQYLLQNSVKQ 457

RESULT 15
ID ROK_RABIT STANDARD; PRT; 463 AA.
AC O19049;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein K (hnRNP K).
GN HNRPK.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP MEDLINE=99207122; PubMed=10101190;
RA Thiele B.J., Berger M., Huth A., Reimann I., Schwarz K., Thiele H.;
RT "Tissue-specific translational regulation of alternative rabbit
RT 15-lipoxygenase mRNAs differing in their 3'-untranslated regions.";
RL Nucleic Acids Res. 27:1828-1836(1999).
CC -1- FUNCTION: ONE OF THE MAJOR PRE-MRNA-BINDING PROTEINS. BINDS
CC TENACIOUSLY TO POLY(C) SEQUENCES. LIKELY TO PLAY A ROLE IN THE
CC NUCLEAR METABOLISM OF HNRNAs, PARTICULARLY FOR PRE-MRNAs THAT
CC CONTAIN CYTIDINE-RICH SEQUENCES. CAN ALSO BIND POLY(C) SINGLE-
CC STRANDED DNA.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR; NUCLEOLAR (BY
CC SIMILARITY).
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ003024; CAA05815.1; .
DR
```

Search completed: April 24, 2003, 16:13:28
Job time : 16 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 ; Search time 40 Seconds
(without alignments)
1928.803 Million cell updates/sec

Title: US-09-897-778-176
Perfect score: 2956
Sequence: 1 MKLYIGLSNAAPSDLES.....VKHQOQKALQSPQSRRK 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

| Database : | | | |
|--------------------|--|--|--|
| A_Geneseq_101002.* | | | |
| 1: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.* | | |
| 2: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.* | | |
| 3: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.* | | |
| 4: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.* | | |
| 5: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.* | | |
| 6: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.* | | |
| 7: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.* | | |
| 8: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.* | | |
| 9: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.* | | |
| 10: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.* | | |
| 11: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.* | | |
| 12: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.* | | |
| 13: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.* | | |
| 14: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.* | | |
| 15: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.* | | |
| 16: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.* | | |
| 17: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.* | | |
| 18: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.* | | |
| 19: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.* | | |
| 20: | /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.* | | |
| 21: | /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.* | | |
| 22: | /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.* | | |
| 23: | /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.* | | |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 2956 | 100.0 | 579 | 21 AAB11328 | Human lung cancer- |
| 2 | 2956 | 100.0 | 579 | 23 ABB74960 | Human lung tumor |
| 3 | 2943 | 99.6 | 579 | 21 AAB11365 | Human lung cancer |
| 4 | 2943 | 99.6 | 579 | 23 ABB74997 | Human lung tumor |
| 5 | 2943 | 99.6 | 579 | 23 ABB75053 | Human lung tumor |
| 6 | 2943 | 99.6 | 579 | 23 ABB75054 | Human lung tumor |
| 7 | 2938 | 99.4 | 586 | 23 ABB75048 | Human lung tumor |
| 8 | 2814 | 95.2 | 619 | 22 ABB21963 | Novel human diagno |
| 9 | 2733 | 92.5 | 583 | 22 ABB12592 | Novel human diagno |
| 10 | 2190 | 74.1 | 577 | 20 AAY30649 | A murine c-myc cod |

| | | | | | | |
|----|--------|------|------|----|----------|---------------------|
| 11 | 1950.5 | 66.0 | 614 | 22 | ABG06794 | Novel human, diagno |
| 12 | 1946.5 | 65.8 | 620 | 22 | AAU16163 | Human novel, secret |
| 13 | 1868 | 63.2 | 594 | 22 | ABG06795 | Novel human, diagno |
| 14 | 1232 | 41.7 | 319 | 22 | AAW93826 | Human polypeptide, |
| 15 | 949 | 32.1 | 261 | 22 | AAU16161 | Human novel, secret |
| 16 | 919 | 31.1 | 250 | 22 | AAU16579 | Human novel, secret |
| 17 | 835.5 | 28.3 | 558 | 22 | ABW58367 | Drosophila melanog |
| 18 | 776.5 | 26.3 | 209 | 22 | ABG21961 | Novel human, diagno |
| 19 | 775.5 | 26.2 | 266 | 22 | ABG12593 | Human novel, diagno |
| 20 | 626 | 21.2 | 171 | 22 | ABG16166 | Human novel, secret |
| 21 | 551 | 18.6 | 148 | 22 | ABG21962 | Novel human, diagno |
| 22 | 527 | 17.8 | 192 | 22 | AAU16164 | Human bone marrow |
| 23 | 474 | 16.0 | 93 | 22 | AAW78238 | Peptide, #12538 enc |
| 24 | 474 | 16.0 | 93 | 22 | AAW38501 | Human novel, secret |
| 25 | 402 | 13.6 | 171 | 22 | AAU16583 | Human novel, secret |
| 26 | 341 | 11.5 | 97 | 21 | AAW03261 | Human secreted pro |
| 27 | 241 | 8.2 | 644 | 21 | AAW58813 | Human c-myc, far up |
| 28 | 231.5 | 7.8 | 313 | 21 | AAW08900 | Arabidopsis, thalia |
| 29 | 229.5 | 7.8 | 590 | 15 | AAW58816 | Human c-myc, far up |
| 30 | 229 | 7.7 | 386 | 22 | ABW63279 | Drosophila melanog |
| 31 | 228.5 | 7.7 | 643 | 15 | AAW58814 | Human c-myc, far up |
| 32 | 227 | 7.7 | 45 | 22 | ABW43427 | Peptide, #10933 enc |
| 33 | 227 | 7.7 | 45 | 22 | AAW64356 | Human brain expres |
| 34 | 227 | 7.7 | 45 | 22 | AAW37315 | Peptide, #11352 enc |
| 35 | 227 | 7.7 | 530 | 13 | AAW28888 | RI paraneoplastic |
| 36 | 226 | 7.6 | 1268 | 22 | ABW61314 | Drosophila melanog |
| 37 | 226 | 7.6 | 1320 | 23 | AAW47603 | Drosophila cell cy |
| 38 | 218 | 7.4 | 507 | 22 | ABG23242 | Novel human, diagno |
| 39 | 217.5 | 7.4 | 644 | 21 | AAW39346 | Arabidopsis, thalia |
| 40 | 213.5 | 7.2 | 403 | 21 | AAW42547 | Human ORF2211 |
| 41 | 213.5 | 7.2 | 403 | 22 | AAW94825 | Human protein sequ |
| 42 | 207.5 | 7.0 | 1310 | 22 | AAU32365 | Novel human secret |
| 43 | 204.5 | 6.9 | 781 | 21 | AAW31213 | Arabidopsis thalia |
| 44 | 204.5 | 6.9 | 803 | 21 | AAW31212 | Arabidopsis thalia |
| 45 | 204.5 | 6.9 | 846 | 21 | AAW31211 | Arabidopsis thalia |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AA11328 | |
| ID | AA11328 standard; Protein; 579 AA. |
| XX | |
| XX | AA11328; |
| XX | |
| XX | |
| DT | 21-FEB-2001 (first entry) |
| XX | |
| DE | Human lung cancer-associated protein L5235. |
| XX | |
| XX | |
| KW | Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic; |
| KW | vaccine; detection. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200061612-A2. |
| XX | |
| PD | 19-OCT-2000. |
| XX | |
| XX | |
| PF | 03-APR-2000; 2000WO-US08896. |
| XX | |
| PR | 02-APR-1999; 99US-0285479. |
| PR | 17-DEC-1999; 99US-0466396. |
| PR | 30-DEC-1999; 99US-0476496. |
| PR | 10-JAN-2000; 2000US-0480884. |
| PR | 22-FEB-2000; 2000US-0510376. |
| XX | |
| XX | (CORI-) CORIXA CORP. |
| PA | |
| PI | Wang T, Fan L; |
| PI | |
| DR | WPI: 2000-628399/60. |
| DR | N-PSDB; AAC65900. |

XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
 PT protein is used for detecting and monitoring progression of lung cancer
 PT in a patient
 XX
 PS Claim 3; Page 186-188; 261pp; English.
 XX
 CC This invention describes a novel isolated polypeptide (I) which
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
 CC which have cytostatic activity. The polypeptides and polynucleotides are
 CC used in compositions and vaccines to inhibit the development of cancer,
 CC especially lung cancer, in a patient. Methods described in the invention
 CC can be used to monitor the progression of a cancer by carrying out the
 CC detection at subsequent time points and comparing the results from the
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting
 CC cells expressing P2 and then administered to the patient to inhibit
 CC development of cancer.
 XX
 SQ Sequence 579 AA;

Query Match 100.0%; Score 2956; DB 21; Length 579;
 Best Local Similarity 100.0%; Pred. No. 9.7e-246;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKLYIGNLSENAAPSDLESIFDKAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60
 Db 1 MNKLYIGNLSENAAPSDLESIFDKAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60

QY 61 IELHGKPIEVEHSPVKRQRIRKQIRNIPPHLQWELVDSLLVQYGVVSECEQVNTDSETA 120
 Db 61 IELHGKPIEVEHSPVKRQRIRKQIRNIPPHLQWELVDSLLVQYGVVSECEQVNTDSETA 120

QY 121 VNVVYSSKQDQARQALDKLNGFLENFTLVKAVIPDEMAAQQNPLOQPRRGLGQSS 180
 Db 121 VNVVYSSKQDQARQALDKLNGFLENFTLVKAVIPDEMAAQQNPLOQPRRGLGQSS 180

QY 181 ROGSPGSVKQKPCDPLRLVPTQFVGALIGKEGATIRNITKQTSKIDVHRKENAGAA 240
 Db 181 ROGSPGSVKQKPCDPLRLVPTQFVGALIGKEGATIRNITKQTSKIDVHRKENAGAA 240

QY 241 EKSITILSTPEGTSAAKSSILEIMHKEAQDIKFTPEEIPKLILAHNNFVGRGLGKEGRNLK 300
 Db 241 EKSITILSTPEGTSAAKSSILEIMHKEAQDIKFTPEEIPKLILAHNNFVGRGLGKEGRNLK 300

QY 301 KIEQDQDTKITISPLQELTLYNPRTITVKGNVETCAKAEEMKIRESYENDIASMNL 360
 Db 301 KIEQDQDTKITISPLQELTLYNPRTITVKGNVETCAKAEEMKIRESYENDIASMNL 360

QY 361 QAHILPGLNLNGLFPPTSGMPPPTSGPPSANTPPYQFQESQETVHOFIPALSVGAI 420
 Db 361 QAHILPGLNLNGLFPPTSGMPPPTSGPPSANTPPYQFQESQETVHOFIPALSVGAI 420

QY 421 IKQGOHILKQSRFAGASTKIAPADPAKVRWVITITGPPPEAOFKAQGRYKIKENFV 480
 Db 421 IKQGOHILKQSRFAGASTKIAPADPAKVRWVITITGPPPEAOFKAQGRYKIKENFV 480

QY 481 SPKEEVKLEAHIRVFSFAGRVIKGGKTVNELQNLSSAEVVVPRDQTPDENDQVVKIT 540
 Db 481 SPKEEVKLEAHIRVFSFAGRVIKGGKTVNELQNLSSAEVVVPRDQTPDENDQVVKIT 540

QY 541 GHFYACQVAQRKIQELTQVKHQHQQKALQSGPPQSRRK 579
 Db 541 GHFYACQVAQRKIQELTQVKHQHQQKALQSGPPQSRRK 579

RESULT 2
 ID ABB74960
 XX ABB74960 standard; Protein; 579 AA.
 AC ABB74960;
 XX
 DT 01-MAY-2002 (first entry)

XX Human lung tumour L523S protein sequence SEQ ID NO:176.
 DE Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 KW immune response.
 XX Homo sapiens.
 OS
 XX WO200200174-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 28-JUN-2001; 2001WO-US21065.
 PF
 XX 28-JUN-2000; 2000US-0606421.
 PR 02-AUG-2000; 2000US-0630940.
 PR 21-AUG-2000; 2000US-0643597.
 PR 15-SEP-2000; 2000US-0662786.
 PR 09-OCT-2000; 2000US-0685696.
 PR 12-DEC-2000; 2000US-0735705.
 PR 07-MAY-2001; 2001US-0850716.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
 PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
 XX WPI: 2002-090513/12.
 DR N-PSDB; ABL49119.
 XX Polynucleotides encoding lung tumor polypeptides, useful for treating
 PT lung cancer or stimulating an immune response -
 PS Example 2; Page 267-268; 374pp; English.
 XX The present invention describes human lung tumour proteins. Human lung
 CC tumour proteins and polynucleotides have cytostatic and immunostimulant
 CC activities, and can be used in vaccine production. Compositions
 CC comprising the lung tumour proteins, polynucleotides, antibodies,
 CC fusion proteins, T cell populations, or antigen presenting cells that
 CC express the lung tumour proteins are useful for treating lung cancer or
 CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 CC ABB75070 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 579 AA;

Query Match 100.0%; Score 2956; DB 23; Length 579;
 Best Local Similarity 100.0%; Pred. No. 9.7e-246;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKLYIGNLSENAAPSDLESIFDKAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60
 Db 1 MNKLYIGNLSENAAPSDLESIFDKAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60

QY 61 IELHGKPIEVEHSPVKRQRIRKQIRNIPPHLQWELVDSLLVQYGVVSECEQVNTDSETA 120
 Db 61 IELHGKPIEVEHSPVKRQRIRKQIRNIPPHLQWELVDSLLVQYGVVSECEQVNTDSETA 120

QY 121 VNVVYSSKQDQARQALDKLNGFLENFTLVKAVIPDEMAAQQNPLOQPRRGLGQSS 180
 Db 121 VNVVYSSKQDQARQALDKLNGFLENFTLVKAVIPDEMAAQQNPLOQPRRGLGQSS 180

QY 181 ROGSPGSVKQKPCDPLRLVPTQFVGALIGKEGATIRNITKQTSKIDVHRKENAGAA 240
 Db 181 ROGSPGSVKQKPCDPLRLVPTQFVGALIGKEGATIRNITKQTSKIDVHRKENAGAA 240

QY 241 EKSITILSTPEGTSAAKSSILEIMHKEAQDIKFTPEEIPKLILAHNNFVGRGLGKEGRNLK 300
 Db 241 EKSITILSTPEGTSAAKSSILEIMHKEAQDIKFTPEEIPKLILAHNNFVGRGLGKEGRNLK 300

QY 301 KIEQDQDTKITISPLQELTLYNPRTITVKGNVETCAKAEEMKIRESYENDIASMNL 360

||||| 301 KIEQDTDKITISPLQELTLYNPERTITVKGNETCAKAEIEIMKKIRSEYENDIASMNL 360
Db
QY 361 QAHLIPGLNIALGLFPPTSGMPPPTSGPPSAMTPPYQFQSETEVTHQIPALSVCAL 420
Db 361 QAHLIPGLNIALGLFPPTSGMPPPTSGPPSAMTPPYQFQSETEVTHQIPALSVCAL 420
QY 421 IGKQGHILKQLSRFAGASIKIAPAEAPDAKVRWVIITGPPPAQKAGRIYKIKKEENFV 480
Db 421 IGKQGHILKQLSRFAGASIKIAPAEAPDAKVRWVIITGPPPAQKAGRIYKIKKEENFV 480
QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
QY 541 GHFYACQVAQRKIQIBLTQVKHQHQQKALQSGPPQSRKK 579
Db 541 GHFYACQVAQRKIQIBLTQVKHQHQQKALQSGPPQSRKK 579
RESULT 3
AABL1365
ID AABL1365 standard; Protein; 579 AA.
AC AABL1365;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human lung cancer associated antigen L523S.
XX
KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection.
XX
OS Homo sapiens.
XX
PN WO200061612-A2.
XX
PD 19-OCT-2000.
XX
PF 03-APR-2000; 2000WO-US08896.
XX
PR 02-APR-1999; 99US-0285479.
PR 17-DEC-1999; 99US-0466396.
PR 30-DEC-1999; 99US-0476496.
PR 10-JAN-2000; 2000US-0480884.
PR 22-FEB-2000; 2000US-0510376.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L;
XX
DR WPI; 2000-628399/60.
DR N-PSDB; AAC66035.
XX
PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT protein is used for detecting and monitoring progression of lung cancer
PT in a patient -
XX
PS Claim 3; Page 259-261; 261pp; English.
XX
CC This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer.
XX
SQ Sequence 579 AA;

Query Match 99.6%; Score 2943; DB 21; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.3e-244;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MNKLYIGNLSAAPSDESIFKDAKIPVSGPFLVKTGYAFVDCPDDESWALKAIEALSCK 60
Db 1 MNKLYIGNLSAAPSDESIFKDAKIPVSGPFLVKTGYAFVDCPDDESWALKAIEALSCK 60
QY 61 IELHGKPIEVHSPKQRIRKIQIRNIPPHLOWEVLDSLIVQYGVVSECEQVNTDSETA 120
Db 61 IELHGKPIEVHSPKQRIRKIQIRNIPPHLOWEVLDSLIVQYGVVSECEQVNTDSETA 120
QY 121 VVNTYSSKQDQARQALDKLNGFOLENTLKVAYIPDMAAQNPLOQPRGRGLGQSS 180
Db 121 VVNTYSSKQDQARQALDKLNGFOLENTLKVAYIPDMAAQNPLOQPRGRGLGQSS 180
QY 181 ROGSPGSVSKQKPCDPLRLLLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Db 181 ROGSPGSVSKQKPCDPLRLLLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
QY 241 EKSITILSTPEGTSACKSILEIMHKEAQDIKETEIPILKILAHNNFVGRGLIGKEGRNLK 300
Db 241 EKSITILSTPEGTSACKSILEIMHKEAQDIKETEIPILKILAHNNFVGRGLIGKEGRNLK 300
QY 301 KIEQDQTKITISPLQELTLYNPERTITVKGNETCAKAEIEIMKKIRSEYENDIASMNL 360
Db 301 KIEQDQTKITISPLQELTLYNPERTITVKGNETCAKAEIEIMKKIRSEYENDIASMNL 360
QY 361 QAHLIPGLNIALGLFPPTSGMPPPTSGPPSAMTPPYQFQSETEVTHQIPALSVCAL 420
Db 361 QAHLIPGLNIALGLFPPTSGMPPPTSGPPSAMTPPYQFQSETEVTHQIPALSVCAL 420
QY 421 IGKQGHILKQLSRFAGASIKIAPAEAPDAKVRWVIITGPPPAQKAGRIYKIKKEENFV 480
Db 421 IGKQGHILKQLSRFAGASIKIAPAEAPDAKVRWVIITGPPPAQKAGRIYKIKKEENFV 480
QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
QY 541 GHFYACQVAQRKIQIBLTQVKHQHQQKALQSGPPQSRKK 579
Db 541 GHFYACQVAQRKIQIBLTQVKHQHQQKALQSGPPQSRKK 579
RESULT 4
ABB74997
ID ABB74997 standard; Protein; 579 AA.
XX
AC ABB74997;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human lung tumour L523S protein sequence SEQ ID NO:348.
XX
KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response.
XX
OS Homo sapiens.
XX
PN WO200200174-A2.
XX
PD 03-JAN-2002.
XX
PF 28-JUN-2001; 2001WO-US21065.
XX
PR 28-JUN-2000; 2000US-0606421.
PR 02-AUG-2000; 2000US-0630940.
PR 21-AUG-2000; 2000US-0643597.
PR 15-SEP-2000; 2000US-0662786.
PR 09-OCT-2000; 2000US-0685696.
PR 12-DEC-2000; 2000US-0735705.

```

PR 07-MAY-2001; 2001US-0850716.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Marnierakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX WPI: 2002-090513/12.
DR N-PSDB; ABL49254.
XX Human lung tumour L523S recombinant protein sequence SEQ ID NO:446;
XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
PT lung cancer or stimulating an immune response -
XX Homo sapiens.
XX Example 2; Page 330-332; 374pp; English.
XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies,
XX fusion proteins, T cell populations, or antigen presenting cells that
XX express the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABL74946 to
XX ABL75070 represent sequences used in the exemplification of the present
XX invention.
XX SQ Sequence 579 AA;
Query Match 99.6%; Score 2943; DB 23; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.3e-244;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60
DB 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60
QY 61 IELHGKPIEVESHVPKQRIRKLRIRNIPPHLOWEVLDSLQYGVVSECEQVNTDSETA 120
DB 61 IELHGKPIEVESHVPKQRIRKLRIRNIPPHLOWEVLDSLQYGVVSECEQVNTDSETA 120
QY 121 VNVVYSSKQDQARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLOQPRGRGLGQSGSS 180
DB 121 VNVVYSSKQDQARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLOQPRGRGLGQSGSS 180
QY 181 RQSPGSVSKQKPCDPLRLVLPVTFQVGAIGKEGATIRNITKQTSKIDVHRKENAGAA 240
DB 181 RQSPGSVSKQKPCDPLRLVLPVTFQVGAIGKEGATIRNITKQTSKIDVHRKENAGAA 240
QY 241 EKSITILSTPEGTSAAKSTILETMHKEADQIKFTEIPLKILAHNNFVGRGLGKEGRNLK 300
DB 241 EKSITILSTPEGTSAAKSTILETMHKEADQIKFTEIPLKILAHNNFVGRGLGKEGRNLK 300
QY 301 KIEQDQTKITISPLQELTYLNPRTITVKGNVETCAKAEETIMKIRESYENDIASMNL 360
DB 301 KIEQDQTKITISPLQELTYLNPRTITVKGNVETCAKAEETIMKIRESYENDIASMNL 360
QY 361 QAHILPCLNALGLPPTSGMPPPTSGPPSAMTPPYPOFQSETEVTHQFIPALSVGAI 420
DB 361 QAHILPCLNALGLPPTSGMPPPTSGPPSAMTPPYPOFQSETEVTHQFIPALSVGAI 420
QY 421 IGRQGHQIKQLSFAGASIKIAPAEAPDAKVRMVIITGPPPEAQFKAQRIYKIKIENFV 480
DB 421 IGRQGHQIKQLSFAGASIKIAPAEAPDAKVRMVIITGPPPEAQFKAQRIYKIKIENFV 480
QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTVNELQNLSSAEVVVPRDQTPDENDQVVKIT 540
DB 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTVNELQNLSSAEVVVPRDQTPDENDQVVKIT 540
QY 541 GHFYACQVADRKIQEILTQVQHQQOKALQSGPPQSRK 579
DB 541 GHFYACQVADRKIQEILTQVQHQQOKALQSGPPQSRK 579

```

```

RESULT 5
ABB75053
XX ABB75053 standard; Protein: 579 AA.
XX AC ABB75053;
XX DT 01-MAY-2002 (first entry)
XX DE Human lung tumour L523S recombinant protein sequence SEQ ID NO:446;
XX KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX OS immune response.
XX OS Homo sapiens.
XX PN WO2002001174-A2.
XX PD 03-JAN-2002.
XX PF 28-JUN-2001; 2001WO-US21065.
XX PR 28-JUN-2000; 2000US-0606421.
XX PR 02-AUG-2000; 2000US-0630940.
XX PR 21-AUG-2000; 2000US-0643597.
XX PR 15-SEP-2000; 2000US-0662786.
XX PR 09-OCT-2000; 2000US-0685696.
XX PR 12-DEC-2000; 2000US-0735705.
XX PR 07-MAY-2001; 2001US-0850716.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Marnierakis M, Fanger GR;
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX WPI: 2002-090513/12.
XX DR N-PSDB; ABL49297.
XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response -
XX Claim 2; Page 365-367; 374pp; English.
XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies,
XX fusion proteins, T cell populations, or antigen presenting cells that
XX express the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABL74946 to
XX ABL75070 represent sequences used in the exemplification of the present
XX invention.
XX SQ Sequence 579 AA;
Query Match 99.6%; Score 2943; DB 23; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.3e-244;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60
DB 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSCK 60
QY 61 IELHGKPIEVESHVPKQRIRKLRIRNIPPHLOWEVLDSLQYGVVSECEQVNTDSETA 120
DB 61 IELHGKPIEVESHVPKQRIRKLRIRNIPPHLOWEVLDSLQYGVVSECEQVNTDSETA 120
QY 121 VNVVYSSKQDQARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLOQPRGRGLGQSGSS 180
DB 121 VNVVYSSKQDQARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLOQPRGRGLGQSGSS 180
QY 181 RQSPGSVSKQKPCDPLRLVLPVTFQVGAIGKEGATIRNITKQTSKIDVHRKENAGAA 240

```


28-JUN-2001; 2001WO-US21065.
28-JUN-2000; 200US-0606421.
02-AUG-2000; 200US-0630940.
21-AUG-2000; 200US-0643597.
15-SEP-2000; 200US-0662786.
09-OCT-2000; 200US-0685696.
12-DEC-2000; 200US-0735705.
07-MAY-2001; 2001US-0850716.
(CORI-) CORIXA CORP.
Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
Vedvick TS, Carter D, Watanabe Y, Peckham DW;
WPI: 2002-0905113/12.
DR N-PSDB; ABL49283.
XX Polynucleotides encoding lung tumor polypeptides, useful for treating
PT lung cancer or stimulating an immune response -
XX
XX
PS Claim 2; Page 354-355; 374pp; English.
XX
XX The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies,
CC fusion proteins, T cell populations, or antigen presenting cells that
CC express the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL48959 to ABL49300 and ABL4946 to
CC ABL75070 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 586 AA;
Query Match 99.4%; Score 2938; DB 23; Length 586;
Best Local Similarity 99.7%; Pred. No. 3.5e-244;
Matches 576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 NKLYIGNLSNAAPSLSEIFKDAKIPVSGPLVKTGFAVDCPDSEWALKAIKALSGKI 61
DB 9 NKLYIGNLSNAAPSLSEIFKDAKIPVSGPLVKTGFAVDCPDSEWALKAIKALSGKI 68
QY 62 ELHGKPIEVHSPKRRQIRKLIQIRNIPPHLQWEVLDSLLVQYGVESCEQVNTDSETAV 121
DB 69 ELHGKPIEVHSPKRRQIRKLIQIRNIPPHLQWEVLDSLLVQYGVESCEQVNTDSETAV 128
QY 122 VNYTYSKQDQALDKLNGFQLENFTLKVAYIPDEMAAQQNPLOQPRRGLGQSSR 181
DB 129 VNYTYSKQDQALDKLNGFQLENFTLKVAYIPDETAQQNPLOQPRRGLGQSSR 188
QY 182 QSGPGSVSKQKCDLPLRLVPTQFVGLIGKEGATIRNITKQTSKIDVHRKENAGAAE 241
DB 189 QSGPGSVSKQKCDLPLRLVPTQFVGLIGKEGATIRNITKQTSKIDVHRKENAGAAE 248
QY 242 KSITILSTPGTSAACKSILEIMHKAQDIKFTTEIPKILAHNFVGLIGKEGRNLKK 301
DB 249 KSITILSTPGTSAACKSILEIMHKAQDIKFTTEIPKILAHNFVGLIGKEGRNLKK 308
QY 302 IEQDTDTKIPISPLQELTYLNPERTITVKGNETCAKAEETMKKIRSYENDIASMLQ 361
DB 309 IEQDTDTKIPISPLQELTYLNPERTITVKGNETCAKAEETMKKIRSYENDIASMLQ 368
QY 362 AHIPLGLNALGLFPPTSCMPPTSGPPSAMTPPYQEQSETETVHQFIPALSVGAI 421
DB 369 AHIPLGLNALGLFPPTSCMPPTSGPPSAMTPPYQEQSETETVHLFIPALSVGAI 428
QY 422 GKQGHITKQLSRFAGASIKITAPAEADKVRVVIITGPPPEAQFKAQGRLYGKIKENFVS 481
DB 429 GKQGHITKQLSRFAGASIKITAPAEADKVRVVIITGPPPEAQFKAQGRLYGKIKENFVS 488
QY 482 PKEEVKLEAHIRVPSFAAGRVICKGKGTYNLQNLSSAEVVPDQTPDENDQVVKITG 541

Db 489 PKEEVKLEAHIRVPSFAAGRVICKGKGTYNLQNLSSAEVVPDQTPDENDQVVKITG 548
QY 542 HFYACQVAQRKIQEILITQVKHQHQKALOSGPPQSRRK 579
Db 549 HFYACQVAQRKIQEILITQVKHQHQKALOSGPPQSRRK 586
RESULT 8
ABG21963
ID ABG21963 standard; Protein; 619 AA.
XX
XX AC ABG21963;
XX
XX DT 18-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #21954.
XX
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder;
KW Homo sapiens.
OS
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-05A0217.
XX
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX
XX DR WPI: 2001-639362/73.
XX
XX N-PSDB; AAS86150.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX PS Claim 20; SEQ ID No 52322; 103pp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 619 AA;
Query Match 95.2%; Score 2814; DB 22;
Best Local Similarity 95.0%; Pred. No. 1.9e-233;
Matches 534; Conservative 8; Mismatches 17; Indels 4; Gaps 1;

QY 1 MNKLYIGNLSENAPDSLEIFKDAKIPVSGPFLVKTGYAVVDCPDSEWALKAEALSCK 60
 DB 37 MNKLYIGNLSENAPDSLEIFKDAKIPVSGPFLVKTGYAVVDCPDSEWALKAEALSCK 96
 QY 61 IELHGKPIEVSHSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 120
 DB 97 IELHGKPIEVSHSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 156
 QY 121 VVNVITYSSKDQARQ----LDKLNGLFQLENFTLVKAYIPDSEMAQQNPLOQPRGRGLQ 176
 DB 157 VVNVITYSSKDQARQALDKLKLNGFQLENFTLVKAYIPDSEMAQQNPLOQPRGRGLQ 216
 QY 177 RGSSRQSGSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 236
 DB 217 RGSSRQSGSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 276
 QY 237 AGAAEKSTILSTPEGTSAAKSIILEIMHKEAQDIKFTTEIPLKILAHNNFVGRGLIGK 296
 DB 277 AGAAEKSTILSTPEGTSAAKSIILEIMHKEAQDIKFTTEIPLKILAHNNFVGRGLIGK 336
 QY 297 RNLKKIEQDQTKITISPLQELTYNPRTITVKGNETCAKAEEMKKIRESYENDIA 356
 DB 337 RNLKKIEQDQTKITISPLQELTYNPRTITVKGNETCAKAEEMKKIRESYENDIA 396
 QY 357 SNLQAHILPGLNGLGFLPPTSGMPPPTSGPPSAMTPPYQPEQSEETVHOFIPALS 416
 DB 397 SNLQAHILPGLNGLGFLPPTSGMPPPTSGPPSAMTPPYQPEQSEETVHOFIPALS 456
 QY 417 VCAIIGKOGQHIKOLSRFAGASIKIAPADAPDAKVRVMTITGPPEAQFKAQGRIGYKKE 476
 DB 457 VRALISKOGQHIKOLSRFAGASIKIAPADAPDAKVRVMTITGPPEAQFKAQGRIGYKKE 516
 QY 477 ENFVSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVPDQTPDENDQV 536
 DB 517 ENFVSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVPDQTPDENDQV 576
 QY 537 VKITGHFYACQVQAKRKIOEILTOVKHQQKALQSGPQSRK 579
 DB 577 VKITGHFYACQVQAKRKIOEILTOVKHQQKALQSGPQSRK 619

RESULT 9
 ABG12592
 ID ABG12592 standard; Protein; 583 AA.
 AC ABG12592;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #12583.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO2001175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YF;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS76779.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX Claim 20; SEQ ID NO 42951; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 583 AA;
 SQ

Query Match 92.5%; Score 2733; DB 22; Length 583;
 Best Local Similarity 92.8%; Pred. NO. 1.6e-226;
 Matches 541; Conservative 6; Mismatches 32; Indels 4; Gaps 1;

QY 1 MNKLYIGNLSENAPDSLEIFKDAKIPVSGPFLVKTGYAVVDCPDSEWALKAEALSCK 60
 DB 1 MNKLYIGNLSENAPDSLEIFKDAKIPVSGPFLVKTGYAVVDCPDSEWALKAEALSCK 60
 QY 61 IELHGKPIEVSHSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 120
 DB 61 IELHGKPIEVSHSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 120
 QY 121 VVNVITYSSKDQARQ----LDKLNGLFQLENFTLVKAYIPDSEMAQQNPLOQPRGRGLQ 176
 DB 121 VVNVITYSSKDQARQALDKLKLNGFQLENFTLVKAYIPDSEMAQQNPLOQPRGRGLQ 180
 QY 177 RGSSRQSGSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 236
 DB 181 RGSSRQSGSVKQKQKCDLPLRLVPTQFVGAIIIGKAGATIRNITKOTSKIDVHRKEN 240
 QY 237 AGAAEKSTILSTPEGTSAAKSIILEIMHKEAQDIKFTTEIPLKILAHNNFVGRGLIGK 296
 DB 241 AGAAEKSTILSTPEGTSAAKSIILEIMHKEAQDIKFTTEIPLKILAHNNFVGRGLIGK 300
 QY 297 RNLKKIEQDQTKITISPLQELTYNPRTITVKGNETCAKAEEMKKIRESYENDIA 356
 DB 301 RNLKKIEQDQTKITISPLQELTYNPRTITVKGNETCAKAEEMKKIRESYENDIA 360
 QY 357 SNLQAHILPGLNGLGFLPPTSGMPPPTSGPPSAMTPPYQPEQSEETVHOFIPALS 416
 DB 361 SNLQAHILPGLNGLGFLPPTSGMPPPTSGPPSAMTPPYQPEQSEETVHOFIPALS 420
 QY 417 VCAIIGKOGQHIKOLSRFAGASIKIAPADAPDAKVRVMTITGPPEAQFKAQGRIGYKKE 476
 DB 421 VRALISKOGQHIKOLSRFAGASIKIAPADAPDAKVRVMTITGPPEAQFKAQGRIGYKKE 480
 QY 477 ENFVSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVPDQTPDENDQV 536
 DB 481 ENFVSPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVPDQTPDENDQV 540
 QY 537 VKITGHFYACQVQAKRKIOEILTOVKHQQKALQSGPQSRK 579

Db 541 VKITGHFYACVAQRKIOEILTVQKHQOQKALQSGPPQSRK 583

RESULT 10

AY30649 standard; Protein: 577 AA.

XX AY30649;

XX 17-NOV-1999 (first entry)

XX A murine c-myc coding region determinant binding protein.

XX c-myc coding region determinant binding protein; CRD-BP; tumor;

XX c-myc; endonucleolytic attack; half-life; breast cancer; colon cancer;

XX pancreatic cancer.

XX Mus musculus.

XX WO9446594-A2.

XX 16-SEP-1999.

XX 05-MAR-1999; 99WO-US04897.

XX 09-MAR-1998; 98US-0077372.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

XX Ross J;

XX WPI; 1999-551506/46.

XX N-PSDB; AA210617.

XX Diagnosing presence or absence of a tumor in a human by examining c-myc

XX coding region determinant-binding protein

XX Example; Fig 1A-D; 79pp; English.

XX The present sequence represents a murine c-myc coding region determinant

XX binding protein (CRD-BP). The presence or absence of a tumor can be

XX determined by determining the levels of CRD-BP present in the suspect

XX tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack

XX and so prolongs its half-life. The methods are used for diagnosing

XX presence or absence of a tumor in a human, especially breast, colon

XX and pancreatic cancer. They are also used to inhibit cancer cell

XX growth.

XX Sequence 577 AA;

XX Query Match 74.1%; Score 2190; DB:20; Length 577;

XX Best Local Similarity 74.1%; Pred. No. 1e-179;

XX Matches 434; Conservative 62; Mismatches 74; Indels 16; Gaps 8;

Qy 1 MNKLYGNLSENAFSDLESTFKDAIPVSGPFLVGTGAFVDPDPDSWALKAIKALSGK 60

Db 1 MNKLYGNLSENVTFADLEKVFABKHISYSGQFLVKSGYAFVDPDPDHWAKAIETFSKG 60

Qy 61 IELHGRPIEYHSPKQRIKQIRNIPPHLQWELVDSLLVQYGVVSECEQVNTDSETA 120

Db 61 VELQGRLEHSHVSPKQSRKQIRNIPPHLQWELVDSLLVQYGVVSECEQVNTDSETA 120

Qy 121 VVNTYSSKDOARQALDKLNGFLENFLKVAITPDEMAQQNPLOQPGRR-GIGORGS 179

Db 121 VVNTYSSKDOARQALDKLNGFLENFLKVAITPDEMAQQNPLOQPGRR-GIGORGS 176

Qy 180 SROGSP---GSVSKQKCDPLRLVPTQFVGAIGKEGATIRNITKQSKIDVHRKEN 236

Db 177 PROGSPVAAGAPAKQOPVDIPLRLVPTQYVGAIGKEGATIRNITKQSKIDVHRKEN 236

Qy 237 AGAERKSTITLSPETSAAKSIIEIMHKAQDKEITEIPUKIAHNNFVRLIGKEG 296

Db 237 AGAERKSTITLSPETSAAKSIIEIMHKAQDKEITEIPUKIAHNNFVRLIGKEG 296

Qy 297 RNKKIEQDQDTKITISPLQELTLYNPRTITVKGWNETCAKAEIEIMKKIRESYENDIA 356

Db 297 RNKKIEQDQDTKITISPLQELTLYNPRTITVKGWNETCAKAEIEIMKKIRESYENDIA 356

Qy 357 SMVLAHLIPGLNMLALGFPPTSGMPPTSGPPSAMT--PPYPOFEQS-ETETVHOEIP 413

Db 357 AMSLQHLIPGLNMLALGFPPTSGMPPTSGPPSAMT--PPYPOFEQS-ETETVHOEIP 413

Qy 414 ALSVGAIIKQGHQIKQLSREAFAGASIKIAPAEAPDAKVRWVITITGPPPAQKAGRIYK 473

Db 414 AQAAGVGAIIKQGHQIKQLSREAFAGASIKIAPAEAPDAKVRWVITITGPPPAQKAGRIYK 473

Qy 474 IKENFVSPKEVKLEAHIRVPSFAAGRVIGKGTVNLQNLSSAEVVPVPRDQTPDEND 533

Db 474 LKENEFGKEVKLEAHIRVPSFAAGRVIGKGTVNLQNLSSAEVVPVPRDQTPDEND 533

Qy 534 QVVVKITGHFYACVAQRKIOEILTVQKHQOQKALQSGPPQSRK 579

Db 534 QVVVKITGHFYACVAQRKIOEILTVQKHQOQKALQSGPPQSRK 579

RESULT 11

ABG06794

ID ABG06794 standard; Protein: 614 AA.

XX AC ABG06794;

XX DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #6785.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PA 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX N-PSDB; AAS70981.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX Claim 20; SEQ ID No 37153; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 614 AA;

Query Match 66.08; Score 1950.5; DB 22; Length 614;
Best Local Similarity 63.38; Pred. No. 4.8e-159;
Matches 386; Conservative 81; Mismatches 95; Indels 29; Gaps 8;

QY 1 MNKLYIGLSENAAPSDIEFIDAKIPVSGFVLKTYAFVDCPDSEWALKATEALSGK 60
DB 17 MNKLYIGLSENAAPSDIEFIDAKIPVSGFVLKTYAFVDCPDSEWALKATEALSGK 76
QY 61 IELHGKPIEVSHPKRIKRLQIRNIPPHLQWELVDSLLVQGVVSCQVNTDSETA 120
DB 77 VELHGKIMEVDYSVKLRSRKIQIRNIPPHLQWELVDSLLVQGVVSCQVNTDSETA 136
QY 121 VVNTYSSKQARALDKINGFLENFTLVKAYIPDEMAAQNLPQPRGRGLGQSGSS 180
DB 137 VVNTYATREEAKIAEKLSGHQFENYSEKISYIPDEEVSPSPPO--RAQR--GDHSSR 192
QY 181 ROG-SPGVSYSKOKCDLPLRLVPTQFVGAIGKEGATIRNITKQTSKIDVHRKENAGA 239
DB 193 EOGHAPGGTQARQIDFPLRLVPTQFVGAIGKEGATIRNITKQTSKIDVHRKENAGA 252
QY 240 AKSTITLSPEGTSAACKSILEIMHKEAQDILKFTTEIPKLTLANNFVRLIGKEGRNL 299
DB 253 AEKPTIHAPEGTSEACRMILEIMQKEADTKLAEIPLKLTLANNFVRLIGKEGRNL 312
QY 300 KTEQDITITISPLQELTLNPERTITVKNVETCAKAEIEIMKKIRESYENDIASMN 359
DB 313 KKHETGTITISPLQELTLNPERTITVKNVETCAKAEIEIMKKIRESYENDIASMN 372
QY 360 LQHLIPGNLNLALGF-----PPTS--GMPPPT-----SGPPSAMTP-----P 396
DB 373 QOANLIPGNLNLALGF-----PPTS--GMPPPT-----SGPPSAMTP-----P 432
QY 397 YPOFEQ-SETETVQIPALSGAIGKOGHILKOLSRFAGASIKIAPAEAPDAKRVMI 455
DB 433 FPHHYSYDSEIVNLFIPTQVGAIGKKAHILKOLSRFAGASIKIAPAEAPDAKRVMI 492
QY 456 ITGPPEAQFKAQRIYKIKENFVSPEEVKLEAHIRVPSFAAGRVIGKGTVMNELQN 515
DB 493 ITGPPEAQFKAQRIYKIKENFVSPEEVKLEAHIRVPSFAAGRVIGKGTVMNELQN 552
QY 516 LSAEAVVPRDQTPDENDQVVKITGHFYACQVQAQRKIQEILTOVKHQHQO 566
DB 553 LSAEAVVPRDQTPDENDQVVKITGHFYACQVQAQRKIQEILTOVKHQHQO 603

RESULT 12
AAU16163
ID AAU16163 standard; Protein; 620 AA.
XX AC AAU16163;
XX DT 07-NOV-2001 (first entry)
XX DE Human novel secreted protein, Seq ID 1116.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; vulnerable; secreted protein; rheumatoid arthritis; hyperproliferative disorder; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; angiogenesis; nervous system disorder; Alzheimer's disease; infection; ocular disorder;

KW corneal infection; wound healing; epithelial cell proliferation; skin ageing; food additive; preservative; antiproliferative.
XX Homo sapiens.
OS WO200155322-A2.
PN 02-AUG-2001.
PD 17-JAN-2001; 2001WO-US01341.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249298.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-488783/53.
 XX N-PSDB; AAS26150.
 DR New nucleic acid molecules encoding 461 human secreted proteins for
 XX diagnosing, preventing, treating or ameliorating medical conditions and
 XX used as food additives or preservatives -
 PS Claim 11; SEQ ID No 1116; 980pp; English.
 XX The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast of liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention.

Query Match 65.8%; Score 1946.5; DB 22; Length 620;
 Best Local Similarity 65.1%; Pred. No. 1.1e-158;
 Matches 385; Conservative 81; Mismatches 96; Indels 29; Gaps 8;

QY 1 MNKLYIGNLSENAAPSDEISIFDKAPIVSGPFLVGTGYAFVDCPDSEWALKATLALSGK 60
 DB 23 MNKLYIGNLSPAVTADDLRLQGLFGRKLPLAGQVLLKSGYAFVDYDPQNWAIETLSGK 82
 QY 61 IELHGKPIEVEHSVPKQRIRKLIQRNIPPHLQWEVLDLVLQYGVESCEQVNTDSETA 120
 DB 83 VELHGKIMEVDYSVSKLRKQIRNIPPHLQWEVLDLGLLAQYGVENVQVNTDSETA 142
 QY 121 VVNYTSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLQQRGRGLGQRSS 180
 DB 143 VVNYTATREEAKTAMEKLSGHQFENYSFKISYIPDEEVSSEPPQ--RAQR--GDHSSR 198
 QY 181 RQG-SPGSVSKQKPCDPLRLVPTQFVGAIGKCATIRNITKOTQSKIDVHRKENAGA 239
 DB 199 EQGHAPGTSQARQIDPRLILVPTQFVGAIGKGLTKNITKOTQSRVDIHRKNSGA 258
 QY 240 AEKSITILSPGTSAAKSIILEMHKEAODIKFTEIPLKTLAHNPNFVGLIKEGNRL 299
 DB 259 AEKPVTHATPEGTSEACRMILEMQKEADTKLAEIPLKTLAHNPNFVGLIKEGNRL 318
 QY 300 KKIEQDQDTKITISPLQELTYNPRTITVKGNVETPKAKAEIEIMKKITRESYENDIASMN 359
 DB 319 KKIEHETGKTITISSLDQLSIYNPRTITVKGTVACASAEIEIMKKITRESYENDIASMN 378
 QY 360 LQAHLPCLNLNALGLF-----PPTS--GMPPPT-----SGPPSAMTP-----p 396
 DB 379 QQANLIPCLNLNALGIFSTGLSVLSPGAPRGPAPPAAPYHPFTTHSGYFSSLYPHHFGP 438
 QY 397 YPOFEQ-SETETVHQFIPALSVGAIGKQGHKQLSRFAGASIKIAEAPDAKVRVVI 455

Db 439 FPHHSVPDEIVNLFTQAVGAIIGKKGAKHAKOLARFAGASIKIAPAGDPDSRMI 498
 Qy 456 ITGPEAQFKAQRIYKKEENFVSPKEVLEAHIRVFPFAGRVIGKGTNVLQ 515
 Db 499 ITGPEAQFKAQRIYKKEENFVSPKEVLEAHIRVFPFAGRVIGKGTNVLQ 558
 Qy 516 LSSAEVVPDQTPDENDDVVKITGHFVACQVAKQKTOELTVQKHQQ 566
 Db 559 LTSAEVVPDQTPDENDDVVKITGHFVACQVAKQKTOELTVQKHQQ 609
 RESULT 13
 ABG06795
 ID ABG06795 standard; Protein; 594 AA.
 AC ABG06795;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #6786.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-054027.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HISE-1) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS70982.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 37154; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 594 AA;

Query Match 63.2%; Score 1868; DB 22; Length 594;
 Best Local Similarity 64.8%; Pred. No. 5.9e-152;
 Matches 368; Conservative 78; Mismatches 96; Indels 26; Gaps 6;
 Qy 1 MNKLYIGNLSENAAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAIKALSGK 60
 Db 40 MNKLYIGNLSPAVTADLLQFLGDRKLPLAGOVLLKSGYAFVDPQNWAIKAIETLSGK 99
 Qy 61 IELGKGVIEVHSVPKRIKLOIRNIPPHLOWELVDSLLVQVGVESCEOVNTDSFA 120
 Db 100 VELHGKIMEVDYVSKKLSRKIRIIRNPPHLOWELVDSLLVQVGVESCEOVNTDSFA 159
 Qy 121 VVNTYSKDOARQALDKLNGFLENFTLVKVIIPDEMAAQONPQOPRRGRLGQRRS 180
 Db 160 VVNTYATREAEKIAMKLSGHGFENYSFKISYIPDEEVSPPQ--RAQR--GDHSSR 215
 Qy 181 RQG-SPGVSQKQPCDPLRLVPTQFVGAIIIGKEGATIRNITKQTSKIDVHRKENAGA 239
 Db 216 EGGHAPGGTSGAQRIDFPLRLVPTQFVGAIIIGKEGATIRNITKQTSKIDVHRKENAGA 275
 Qy 240 AEKSITILSTPEGTSAAKSIKLEIMHKEAQDIKFTTEIPKILAHNFYGRIGKGRNL 299
 Db 276 AEKPTIHTPEGTSACRMILEIMKQKADETKLAEEIPLKILAHNGLVGRIGKGRNL 335
 Qy 300 KKIEQDITKTIISPLQELTYNPERTITVKNVETCAKAEIEIMKKIRSYENDIASMN 359
 Db 336 KKIEHETGTRITISLQDLISYINPERTITVKNVETCAKAEIEIMKKIRSYENDIASMN 395
 Qy 360 LQAHILPGLNLAGLFPPTSGMPPPTSGPVSAMTTPYPOFEQ-SETETVHQFIPALSVG 418
 Db 396 THS-----GYF-----SSLYPHQFGPPHHHSYPEQEVNLFIPQAVG 435
 Qy 419 ALICKOGHIKQLSRFAGASIKIAPAEAPDAKVRWVITGTPPEAQFKAQRIYGVKKEEN 478
 Db 436 ALICKGAHILKQLARFAGASIKIAPAEAPDAKVRWVITGTPPEAQFKAQRIYGVKKEEN 495
 Qy 479 FVSPKEVLEAHIRVSPFAGRVIGKGTNVLQNLSSAEVVPDQTPDENDDVVK 538
 Db 496 FNPKEVLEAHIRVSPFAGRVIGKGTNVLQNLSSAEVVPDQTPDENDDVVK 555
 Qy 539 ITGHFYACQVAKQRIKQELTVQKHQQ 566
 Db 556 ITGHFYACQVAKQRIKQELTVQKHQQ 583
 RESULT 14
 AAM93826
 ID AAM93826 standard; Protein; 319 AA.
 XX
 AC AAM93826;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3887.
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isoqai T, Hayashi K, Ishii S., Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI: 2001-524235/58.
 DR N-PSDB; AAK94762.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 use in genetic manipulation
 XX
 PS Claim 8; SEQ ID NO 3887; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ Sequence 319 AA;
 Query Match 41.7%; Score 1232; DB 22; Length 319;
 Best Local Similarity 76.5%; Pred. No. 1.5e-97;
 Matches 247; Conservative 32; Mismatches 36; Indels 8; Gaps 5;
 QY 260 ILEIMHKAQDIKFTTEIPLKLAHNNFVGLIGKEGRNLKKIEQDTDKITISPLQELT 319
 DB 2 ILEIMHKAQDKTADDEVPLKLAHNNFVGLIGKEGRNLKKVEQDTETKTISSLDLT 61
 QY 320 LYPERTIVKGNVETCAKAEETMKIRSYENDIASMNLQALHPLGLNLALGLFPPT 379
 DB 62 LYPERTIVKGAIECCRAQEIMKVRAYENDVAMSLQSHLIPGLNLAAVGLFPAS 121
 QY 380 SGMPPTSGPPSMT--PPYQFQGS-ETETVHOFPALSVGAIGKQGHKQLSRFAG 436
 DB 122 SSAPVP--PPSSVTGAAPYSFQAPQEMVQVFIQAQVGAIGKQGHKQLSRFAS 178
 QY 437 ASIKIAPAEAPAKYRWIITGPPPAQFKAQRIYKIKENFVSPKEVKLEAHIRVPS 496
 DB 179 ASIKIAPPETPSKRVWIIITGPPPAQFKAQRIYKIKENFVSPKEVKLETHIRVPA 238
 QY 497 FAGRVIGKGGTVNELQNLSSAEVVPVROTPDENDVYVKITGHFYACVAQKIOEI 556
 DB 239 SAAGRVIGKGGTVNELQNLTAEEVVPVROTPDENDVYVKITGHFYASQMAQRKIRDI 298
 QY 557 LTQVKHQOQKALQSPQSRKK 579
 DB 299 LAQVKQ-QHQKG-QSNQAQARRK 319
 RESULT 15
 ID AAU16161 standard; Protein: 261 AA.
 AC AAU16161;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human novel secreted protein, Seq ID 1114.
 XX
 KW Human: immunosuppressive; antiarthritic; antirheumatic;
 KW cytotatic; cardiant; vasotropic; cerebroprotective; nootropic;
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
 KW vulnerary; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.

XX Homo sapiens.
 OS WO200155322-A2.
 PN
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01341.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0215647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.

26-SEP-2000; 2000US-0235484.
 27-SEP-2000; 2000US-0235834.
 27-SEP-2000; 2000US-0235836.
 29-SEP-2000; 2000US-0236327.
 29-SEP-2000; 2000US-0236367.
 29-SEP-2000; 2000US-0236368.
 29-SEP-2000; 2000US-0236369.
 29-SEP-2000; 2000US-0236370.
 29-SEP-2000; 2000US-0236802.
 02-OCT-2000; 2000US-0237037.
 02-OCT-2000; 2000US-0237038.
 02-OCT-2000; 2000US-0237039.
 02-OCT-2000; 2000US-0237040.
 13-OCT-2000; 2000US-0239935.
 13-OCT-2000; 2000US-0239937.
 20-OCT-2000; 2000US-0240960.
 20-OCT-2000; 2000US-0241221.
 20-OCT-2000; 2000US-0241785.
 20-OCT-2000; 2000US-0241786.
 20-OCT-2000; 2000US-0241787.
 20-OCT-2000; 2000US-0241808.
 20-OCT-2000; 2000US-0241809.
 20-OCT-2000; 2000US-0241826.
 01-NOV-2000; 2000US-0244617.
 08-NOV-2000; 2000US-0244674.
 08-NOV-2000; 2000US-0244675.
 08-NOV-2000; 2000US-0246476.
 08-NOV-2000; 2000US-0246477.
 08-NOV-2000; 2000US-0246478.
 08-NOV-2000; 2000US-0246523.
 08-NOV-2000; 2000US-0246524.
 08-NOV-2000; 2000US-0246525.
 08-NOV-2000; 2000US-0246526.
 08-NOV-2000; 2000US-0246527.
 08-NOV-2000; 2000US-0246528.
 08-NOV-2000; 2000US-0246532.
 08-NOV-2000; 2000US-0246609.
 08-NOV-2000; 2000US-0246610.
 08-NOV-2000; 2000US-0246611.
 08-NOV-2000; 2000US-0246613.
 17-NOV-2000; 2000US-0249207.
 17-NOV-2000; 2000US-0249208.
 17-NOV-2000; 2000US-0249209.
 17-NOV-2000; 2000US-0249210.
 17-NOV-2000; 2000US-0249211.
 17-NOV-2000; 2000US-0249212.
 17-NOV-2000; 2000US-0249213.
 17-NOV-2000; 2000US-0249214.
 17-NOV-2000; 2000US-0249215.
 17-NOV-2000; 2000US-0249216.
 17-NOV-2000; 2000US-0249217.
 17-NOV-2000; 2000US-0249218.
 17-NOV-2000; 2000US-0249244.
 17-NOV-2000; 2000US-0249245.
 17-NOV-2000; 2000US-0249264.
 17-NOV-2000; 2000US-0249265.
 17-NOV-2000; 2000US-0249297.
 17-NOV-2000; 2000US-0249299.
 17-NOV-2000; 2000US-0249300.
 01-DEC-2000; 2000US-0250160.
 01-DEC-2000; 2000US-0250391.
 05-DEC-2000; 2000US-0251030.
 05-DEC-2000; 2000US-0251988.
 05-DEC-2000; 2000US-0256719.
 06-DEC-2000; 2000US-0251479.
 08-DEC-2000; 2000US-0251856.
 08-DEC-2000; 2000US-0251868.
 08-DEC-2000; 2000US-0251869.
 08-DEC-2000; 2000US-0251989.
 08-DEC-2000; 2000US-0251990.
 11-DEC-2000; 2000US-0254097.
 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPT; 2001-488783/53.
 N-PSDB; AAS26148.

New nucleic acid molecules encoding 461 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions and
 used as food additives or preservatives -

Claim 11; SEQ ID No 1114; 980pp; English.

The invention relates to isolated nucleic acid molecules and their
 encoded secreted proteins. The nucleic acids and proteins are used to
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 rabbits, goats, horses, cats, dogs, chickens or sheep. They
 are also used in diagnosing a pathological condition or susceptibility,
 to a pathological condition. Antibodies to the proteins can also
 be used in alleviating symptoms associated with the disorders and in
 diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 immunosorbant assays (ELISA). Disorders which are diagnosed or treated
 include autoimmune diseases e.g. rheumatoid arthritis,
 hyperproliferative disorders e.g. neoplasms of the breast or liver,
 cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 Alzheimer's disease, infections caused by bacteria, viruses and fungi
 and ocular disorders e.g. corneal infection, and many other
 disorders listed in the specification. The polypeptides can also
 be used to aid wound healing and epithelial cell proliferation, to
 prevent skin aging due to sunburn, to maintain organs before
 transplantation, for supporting cell culture of primary tissues, to
 regenerate tissues and in chemotaxis. The polypeptides can also be used
 as a food additive or preservative to increase or decrease storage
 capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 minerals, cofactors and other nutritional components. The present
 sequence represents a novel secreted protein of the invention.

Query Match 32.1%; Score 949; DB 22; Length 261;
 Best Local Similarity 74.5%; Pred. No. 2 6e-73;
 Matches 193; Conservative 26; Mismatches 32; Indels 8; Gaps 5;

Qy 324 ERTITVKGNETCAKAEIEIMKKITREYENDIASMNIQAHLIPGLNLNALGLFPPTSGMP 383
 Db 8 ERTITVKGAIENCCRAEQEIMKKVREAYENDVAAMSLSHLIPGLNLAAVGLFPASSAV 67
 Qy 384 PPTSGPPSANT--PPYPQFEQS-ETETVHQFIPALSVAIGKQGHKQISRFAGASIK 440
 Db 68 PP---PPSSVTGAAPYSFMOAPEQEMVQVFIQAQAVYGAIGKQGHKQISRFASASIK 124
 Qy 441 IAPAEAPDAKVRMVIITGPPEAQFKAQGRYIKTEENFYSPKEEVKLEAHIRVPSFAAG 500
 Db 125 IAPPETPDSKVRMVIITGPPEAQFKAQGRYIKLKEENFEGPKEEVKLETHIRVPSAAG 184
 Qy 501 RVIGKGGKTVELNLSSAEVVPDPTDENDQVVKITGHFYACQVQAKRKIOEILTQV 560
 Db 185 RVIGKGGKTVELNLSSAEVVPDPTDENDQVVKITGHFYACQVQAKRKIRKIDILAOV 244
 Qy 561 KOHQOQKALQSGPPQSRKK 579
 Db 245 KQ-QHQKG-QSNQAQARRK 261

Search completed: April 24, 2003, 16:15:28
 Job time : 43 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 : Search time 22 Seconds
(without alignments)
2108.871 Million cell updates/sec

Title: US-09-897-778-176
Perfect score: 2956
Sequence: 1 MNKLYIGNLSNAAPSDLES.....VKHQOQKALQSGPPQSRK 579

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 2956 | 100.0 | 579 | 10 | US-09-735-705-176 |
| 2 | 2956 | 100.0 | 579 | 10 | US-09-830-716A-176 |
| 3 | 2956 | 100.0 | 579 | 10 | US-09-897-778-176 |
| 4 | 2943 | 93.6 | 579 | 10 | US-09-735-705-348 |
| 5 | 2943 | 93.6 | 579 | 10 | US-09-850-716A-348 |
| 6 | 2943 | 93.6 | 579 | 10 | US-09-897-778-348 |
| 7 | 2943 | 93.6 | 579 | 10 | US-09-897-778-446 |
| 8 | 2943 | 93.6 | 579 | 10 | US-09-897-778-449 |
| 9 | 2938 | 99.4 | 586 | 10 | US-09-850-716A-427 |
| 10 | 2938 | 99.4 | 586 | 10 | US-09-897-778-427 |
| 11 | 2190 | 74.1 | 577 | 10 | US-09-873-637-2 |
| 12 | 1946.5 | 65.8 | 620 | 10 | US-09-764-864-1116 |
| 13 | 949 | 32.1 | 261 | 10 | US-09-764-864-1114 |
| 14 | 919 | 31.1 | 250 | 10 | US-09-764-864-1532 |
| 15 | 626 | 21.2 | 171 | 10 | US-09-764-864-1119 |
| 16 | 527 | 17.8 | 192 | 10 | US-09-764-864-1117 |
| 17 | 474 | 16.0 | 93 | 10 | US-09-864-761-48606 |
| 18 | 402 | 13.6 | 171 | 10 | US-09-764-864-1536 |
| 19 | 238 | 8.1 | 49 | 10 | US-09-873-637-22 |

ALIGNMENTS

RESULT 1
US-09-735-705-176
Sequence 176, Application US/09735705
Patient NO. US20020052329A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY.

FILE REFERENCE: 210121.455C14

CURRENT APPLICATION NUMBER: US/09/735.705

CURRENT FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 419

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 176

LENGTH: 579

TYPE: PRT

ORGANISM: Homo sapiens

US-09-735-705-176

Query Match 100.0% Score 2956; DB 10; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-217;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | |
|----|----|--|
| Qy | 1 | MNKLYIGNLSNAAPSDLESIFKDAKIPVSGFLVKTGYAFVDCPDSEWALKATFALSGK 60 |
| Db | 1 | MNKLYIGNLSNAAPSDLESIFKDAKIPVSGFLVKTGYAFVDCPDSEWALKATFALSGK 60 |
| Qy | 61 | TELHGKPIEVHSHVPRQRIRKLRQIRNPPLHQWEVLDLSLLVOYGVVSCQVNTDSETA 120 |
| Db | 61 | TELHGKPIEVHSHVPRQRIRKLRQIRNPPLHQWEVLDLSLLVOYGVVSCQVNTDSETA 120 |

Qy 121 VNVYSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGQSS 180
Db 121 VNVYSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGQSS 180
Qy 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Db 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Qy 241 EKSITILSTPGTSAACKSILEIMHKEADIKFTEETPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSITILSTPGTSAACKSILEIMHKEADIKFTEETPLKILAHNNFVGRLLIGKEGRNLK 300
Qy 301 KIEQDQTKITISPLQELTYLNPRTITVKGNETCAKAEEMKKIKRESYENDIASNKL 360
Db 301 KIEQDQTKITISPLQELTYLNPRTITVKGNETCAKAEEMKKIKRESYENDIASNKL 360
Qy 361 QAHLPGLNLNALGLFPPTSGMPPPTSGPPSAMPPTPPYQFQSETEVTHQFIPALSVGAI 420
Db 361 QAHLPGLNLNALGLFPPTSGMPPPTSGPPSAMPPTPPYQFQSETEVTHQFIPALSVGAI 420
Qy 421 IGKQGHOKLSRFAGASIKIAPAEAPDAKVRWIIITGPPPAQKAGRIYKIKKEENFV 480
Db 421 IGKQGHOKLSRFAGASIKIAPAEAPDAKVRWIIITGPPPAQKAGRIYKIKKEENFV 480
Qy 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTNVLONLSSAEVVPVPRDQTPDENDQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTNVLONLSSAEVVPVPRDQTPDENDQVVKIT 540
Qy 541 GHFYACQVQAKRIQEIILTVQKHQKQKALQSGPPQSRKK 579
Db 541 GHFYACQVQAKRIQEIILTVQKHQKQKALQSGPPQSRKK 579

RESULT 2

US-09-850-716A-176 Application US/09850716A
; Sequence 176, Application US/09850716A
; Patent No. US20020115139A1

GENERAL INFORMATION:

; APPLICANT: Kaelos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15

; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-850-716A-176

Query Match 100.0%; Score 2956; DB 10; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-217;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSGK 60
Db 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSGK 60
Qy 61 IELHGKPIEVEHSPVKRIRKIQIRNIPPHLQWEVLDLVLQYGVVSECEQVNTDSETA 120
Db 61 IELHGKPIEVEHSPVKRIRKIQIRNIPPHLQWEVLDLVLQYGVVSECEQVNTDSETA 120
Qy 121 VNVYSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGQSS 180
Db 121 VNVYSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGQSS 180
Qy 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Db 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240

Qy 241 EKSITILSTPGTSAACKSILEIMHKEADIKFTEETPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSITILSTPGTSAACKSILEIMHKEADIKFTEETPLKILAHNNFVGRLLIGKEGRNLK 300
Qy 301 KIEQDQTKITISPLQELTYLNPRTITVKGNETCAKAEEMKKIKRESYENDIASNKL 360
Db 301 KIEQDQTKITISPLQELTYLNPRTITVKGNETCAKAEEMKKIKRESYENDIASNKL 360
Qy 361 QAHLPGLNLNALGLFPPTSGMPPPTSGPPSAMPPTPPYQFQSETEVTHQFIPALSVGAI 420
Db 361 QAHLPGLNLNALGLFPPTSGMPPPTSGPPSAMPPTPPYQFQSETEVTHQFIPALSVGAI 420
Qy 421 IGKQGHOKLSRFAGASIKIAPAEAPDAKVRWIIITGPPPAQKAGRIYKIKKEENFV 480
Db 421 IGKQGHOKLSRFAGASIKIAPAEAPDAKVRWIIITGPPPAQKAGRIYKIKKEENFV 480
Qy 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTNVLONLSSAEVVPVPRDQTPDENDQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGTNVLONLSSAEVVPVPRDQTPDENDQVVKIT 540
Qy 541 GHFYACQVQAKRIQEIILTVQKHQKQKALQSGPPQSRKK 579
Db 541 GHFYACQVQAKRIQEIILTVQKHQKQKALQSGPPQSRKK 579

RESULT 3

US-09-897-778-176

; Sequence 176, Application US/09897778

; Patent No. US20020147143A1

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Manerakis, Margarita

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Peckham, David W.

; APPLICANT: Fanger, Nell

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C16

; CURRENT APPLICATION NUMBER: US/09/897,778

; CURRENT FILING DATE: 2001-06-28

; NUMBER OF SEQ ID NOS: 467

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 176

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-897-778-176

Query Match 100.0%; Score 2956; DB 10; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-217;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSGK 60
Db 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAEALSGK 60
Qy 61 IELHGKPIEVEHSPVKRIRKIQIRNIPPHLQWEVLDLVLQYGVVSECEQVNTDSETA 120
Db 61 IELHGKPIEVEHSPVKRIRKIQIRNIPPHLQWEVLDLVLQYGVVSECEQVNTDSETA 120
Qy 121 VNVYSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGQSS 180
Db 121 VNVYSSKQARQALDKLNGFQLENFTLKVAYIPDEMAAQNPLOQPRGRGLGQSS 180
Qy 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Db 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240

```
Qy 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Qy 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Db 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Qy 361 QAHLIPGLNLAGLFPPTSGMPPTSGPPSAMTPPYQFEQSETVTHQFIPALSVGAI 420
Db 361 QAHLIPGLNLAGLFPPTSGMPPTSGPPSAMTPPYQFEQSETVTHQFIPALSVGAI 420
Qy 421 IGKOGHILKOLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAAGRIYGIKEENFV 480
Db 421 IGKOGHILKOLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAAGRIYGIKEENFV 480
Qy 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Qy 541 GHFYACOVADRKTOEILTQVKHOOQKALOSGPPQSRKK 579
Db 541 GHFYACOVADRKTOEILTQVKHOOQKALOSGPPQSRKK 579

RESULT 4
US-09-735-705-348
; Sequence 348, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Banghr, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fang, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-705-348

Query Match 99.6%; Score 2943; DB 10; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.7e-216;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDERSWALKAEALSGK 60
Db 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDERSWALKAEALSGK 60
Qy 61 IELHGKPIEYHVSVPKRQIRKLQIRNIPPHLOWEVLDSLVLQYGVVESCQVNTDSETA 120
Db 61 IELHGKPIEYHVSVPKRQIRKLQIRNIPPHLOWEVLDSLVLQYGVVESCQVNTDSETA 120
Qy 121 VVNYTSSKQDQARQALDKLNGFLENFTLKVAIIPDEMAAQNPLQOPRGRGLGQSGSS 180
Db 121 VVNYTSSKQDQARQALDKLNGFLENFTLKVAIIPDEMAAQNPLQOPRGRGLGQSGSS 180
Qy 181 RQSGPSVSKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKOTQSKIDVHKEKNAGAA 240
Db 181 RQSGPSVSKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKOTQSKIDVHKEKNAGAA 240
Qy 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Qy 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Db 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Qy 361 QAHLIPGLNLAGLFPPTSGMPPTSGPPSAMTPPYQFEQSETVTHQFIPALSVGAI 420
Db 361 QAHLIPGLNLAGLFPPTSGMPPTSGPPSAMTPPYQFEQSETVTHQFIPALSVGAI 420
Qy 421 IGKOGHILKOLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAAGRIYGIKEENFV 480
Db 421 IGKOGHILKOLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAAGRIYGIKEENFV 480
Qy 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Qy 541 GHFYACOVADRKTOEILTQVKHOOQKALOSGPPQSRKK 579
Db 541 GHFYACOVADRKTOEILTQVKHOOQKALOSGPPQSRKK 579
```

```
Db 181 RQSGPSVSKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKOTQSKIDVHKEKNAGAA 240
Qy 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Qy 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Db 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Qy 361 QAHLIPGLNLAGLFPPTSGMPPTSGPPSAMTPPYQFEQSETVTHQFIPALSVGAI 420
Db 361 QAHLIPGLNLAGLFPPTSGMPPTSGPPSAMTPPYQFEQSETVTHQFIPALSVGAI 420
Qy 421 IGKOGHILKOLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAAGRIYGIKEENFV 480
Db 421 IGKOGHILKOLSRFAGASIKIAPAEAPDAKVRMVIITGPPEAQFKAAGRIYGIKEENFV 480
Qy 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGTVNELQNLSSAEVVPVPRDQTPDENQVVKIT 540
Qy 541 GHFYACOVADRKTOEILTQVKHOOQKALOSGPPQSRKK 579
Db 541 GHFYACOVADRKTOEILTQVKHOOQKALOSGPPQSRKK 579

RESULT 5
US-09-850-716A-348
; Sequence 348, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-348

Query Match 99.6%; Score 2943; DB 10; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.7e-216;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDERSWALKAEALSGK 60
Db 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDERSWALKAEALSGK 60
Qy 61 IELHGKPIEYHVSVPKRQIRKLQIRNIPPHLOWEVLDSLVLQYGVVESCQVNTDSETA 120
Db 61 IELHGKPIEYHVSVPKRQIRKLQIRNIPPHLOWEVLDSLVLQYGVVESCQVNTDSETA 120
Qy 121 VVNYTSSKQDQARQALDKLNGFLENFTLKVAIIPDEMAAQNPLQOPRGRGLGQSGSS 180
Db 121 VVNYTSSKQDQARQALDKLNGFLENFTLKVAIIPDEMAAQNPLQOPRGRGLGQSGSS 180
Qy 181 RQSGPSVSKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKOTQSKIDVHKEKNAGAA 240
Db 181 RQSGPSVSKQKPCDPLRLVPTQFVGAIIGKEGATIRNITKOTQSKIDVHKEKNAGAA 240
Qy 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSTITLSTPEGTSAAKSSILEIMHKEAODIKFTTEIPLKILAHNNFVGRLLIGKEGRNLK 300
Qy 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
Db 301 KIEQDITDKTITISPLQELTYLPNTITVKNVETCAKAEIEIMKKITRESYENDIASMNL 360
```

```
Db 301 KIEODTDTKITISPLQELTLYNPRTITVKGNETCAKAEEMKKIRESYENDIASMNL 360
QY 361 QAHLIPLGLNLNAGLFPPTSGMPPPTSGPPSAMTPPYPOEQSETETVHOFIPALSVCAGI 420
Db 361 QAHLIPLGLNLNAGLFPPTSGMPPPTSGPPSAMTPPYPOEQSETETVHOFIPALSVCAGI 420
QY 421 IGKOGQHIKOLSRFAGASTIKIAPAEAPDAKVRWVIITGPPPEAFKAQRIYKIKKEENFV 480
Db 421 IGKOGQHIKOLSRFAGASTIKIAPAEAPDAKVRWVIITGPPPEAFKAQRIYKIKKEENFV 480
QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVPDQTPDENDQVVVKIT 540
Db 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNELQNLSSAEVVPDQTPDENDQVVVKIT 540
QY 541 GHFYACOVAQRKIQEILTQVKGHOQKALQSGPPQSRK 579
Db 541 GHFYACOVAQRKIQEILTQVKGHOQKALQSGPPQSRK 579
```

RESULT 6

```
US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US20020147143A1
```

GENERAL INFORMATION:

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
```

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

```
; FILE REFERENCE: 210121.455C16
```

```
; CURRENT APPLICATION NUMBER: US/09/897,778
```

```
; NUMBER OF SEQ ID NOS: 467
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 348
```

```
; LENGTH: 579
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-897-778-348
```

```
Query Match 99.6%; Score 2943; DB 10; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.7e-216;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MNKLYIGNLSENAAPSDLESIPFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAIKALSCK 60
Db 1 MNKLYIGNLSENAAPSDLESIPFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAIKALSCK 60
QY 61 IELHGKPIEVHSPVKRQIRKLIQIRNIPPHLOWEVLDSLVLQYGVVSECEQVNTDSETA 120
Db 61 IELHGKPIEVHSPVKRQIRKLIQIRNIPPHLOWEVLDSLVLQYGVVSECEQVNTDSETA 120
QY 121 VVWYTSKDKQARQALDKLNGFLENFTLKVAYIPDEMAAQNPLOQPRRGLGQSGSS 180
Db 121 VVWYTSKDKQARQALDKLNGFLENFTLKVAYIPDEMAAQNPLOQPRRGLGQSGSS 180
QY 181 RQSPGVSQKQPCDPLRLLLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Db 181 RQSPGVSQKQPCDPLRLLLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
QY 241 EKSTITLSTPEGTSACKSILIMHKEAQDIKFTFEEIPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSTITLSTPEGTSACKSILIMHKEAQDIKFTFEEIPLKILAHNNFVGRLLIGKEGRNLK 300
QY 301 KIEODTDTKITISPLQELTLYNPRTITVKGNETCAKAEEMKKIRESYENDIASMNL 360
Db 301 KIEODTDTKITISPLQELTLYNPRTITVKGNETCAKAEEMKKIRESYENDIASMNL 360
```

RESULT 7

```
US-09-897-778-446
```

```
; Sequence 446, Application US/09897778
; Patent No. US20020147143A1
```

GENERAL INFORMATION:

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
```

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

```
; FILE REFERENCE: 210121.455C16
```

```
; CURRENT APPLICATION NUMBER: US/09/897,778
```

```
; NUMBER OF SEQ ID NOS: 467
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 446
```

```
; LENGTH: 579
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-897-778-446
```

```
Query Match 99.6%; Score 2943; DB 10; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.7e-216;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MNKLYIGNLSENAAPSDLESIPFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAIKALSCK 60
Db 1 MNKLYIGNLSENAAPSDLESIPFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKAIKALSCK 60
QY 61 IELHGKPIEVHSPVKRQIRKLIQIRNIPPHLOWEVLDSLVLQYGVVSECEQVNTDSETA 120
Db 61 IELHGKPIEVHSPVKRQIRKLIQIRNIPPHLOWEVLDSLVLQYGVVSECEQVNTDSETA 120
QY 121 VVWYTSKDKQARQALDKLNGFLENFTLKVAYIPDEMAAQNPLOQPRRGLGQSGSS 180
Db 121 VVWYTSKDKQARQALDKLNGFLENFTLKVAYIPDEMAAQNPLOQPRRGLGQSGSS 180
QY 181 RQSPGVSQKQPCDPLRLLLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
Db 181 RQSPGVSQKQPCDPLRLLLVPTQFVGAIIGKEGATIRNITKQTSKIDVHRKENAGAA 240
QY 241 EKSTITLSTPEGTSACKSILIMHKEAQDIKFTFEEIPLKILAHNNFVGRLLIGKEGRNLK 300
Db 241 EKSTITLSTPEGTSACKSILIMHKEAQDIKFTFEEIPLKILAHNNFVGRLLIGKEGRNLK 300
QY 301 KIEODTDTKITISPLQELTLYNPRTITVKGNETCAKAEEMKKIRESYENDIASMNL 360
Db 301 KIEODTDTKITISPLQELTLYNPRTITVKGNETCAKAEEMKKIRESYENDIASMNL 360
```

QY 361 QAHLIPGLNGLALGLFPPTSGMPPTSGPPSMTPTTYPQFQSETEVHOFIPALSVGAI 420
DB 361 QAHLIPGLNGLNGLALGLFPPTSGMPPTSGPPSMTPTTYPQFQSETEVHOFIPALSVGAI 420
QY 421 IGKQGHKQLSFRFAGASIKIAPAEAPDAKVRWVITITGPEAKQAGRIYKIKENFV 480
DB 421 IGKQGHKQLSFRFAGASIKIAPAEAPDAKVRWVITITGPEAKQAGRIYKIKENFV 480
QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVVPRDQTPDENDQVVKIT 540
DB 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVVPRDQTPDENDQVVKIT 540
QY 541 GHFYACQVAQRKIQEILTOVKHQHQQKALQSGPPQSRK 579
DB 541 GHFYACQVAQRKIQEILTOVKHQHQQKALQSGPPQSRK 579

RESULT 8
US-09-897-778-449
; Sequence 449, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-449

Query Match 99.6%; Score 2943; DB 10; Length 579;
Best Local Similarity 99.7%; Pred. No. 1.7e-216;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKATEALS 60
DB 1 MNKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKATEALS 60
QY 61 ELHGKPIEVESHVPKQRIKQIRNIPPHLQWVDSLVVQGVVSCQVNTDSETA 120
DB 61 ELHGKPIEVESHVPKQRIKQIRNIPPHLQWVDSLVVQGVVSCQVNTDSETA 120
QY 121 VNYTYSKDOARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLQOPRGRGLGQSS 180
DB 121 VNYTYSKDOARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLQOPRGRGLGQSS 180
QY 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIIGEGATIRNITKQTSKIDVHRKENAGAA 240
DB 181 ROGSPGSVKQKPCDPLRLVPTQFVGAIIIGEGATIRNITKQTSKIDVHRKENAGAA 240
QY 241 EKSTILSTPEGSTSAACKSILEIMHKEAQDIKFTTEIPLKILAHNNFVGRILGEGRNK 300
DB 241 EKSTILSTPEGSTSAACKSILEIMHKEAQDIKFTTEIPLKILAHNNFVGRILGEGRNK 300
QY 301 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 360
DB 301 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 360

QY 361 QAHLIPGLNGLALGLFPPTSGMPPTSGPPSMTPTTYPQFQSETEVHOFIPALSVGAI 420
DB 361 QAHLIPGLNGLNGLALGLFPPTSGMPPTSGPPSMTPTTYPQFQSETEVHOFIPALSVGAI 420
QY 421 IGKQGHKQLSFRFAGASIKIAPAEAPDAKVRWVITITGPEAKQAGRIYKIKENFV 480
DB 421 IGKQGHKQLSFRFAGASIKIAPAEAPDAKVRWVITITGPEAKQAGRIYKIKENFV 480
QY 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVVPRDQTPDENDQVVKIT 540
DB 481 SPKEEVKLEAHIRVPSFAAGRVIGKGGKTVNLQNLSSAEVVVPRDQTPDENDQVVKIT 540
QY 541 GHFYACQVAQRKIQEILTOVKHQHQQKALQSGPPQSRK 579
DB 541 GHFYACQVAQRKIQEILTOVKHQHQQKALQSGPPQSRK 579

RESULT 9
US-09-850-716A-427
; Sequence 427, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Michael D.
; APPLICANT: Retter, Mac W.
; APPLICANT: Retter, Mac W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 427
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-427

Query Match 99.4%; Score 2938; DB 10; Length 586;
Best Local Similarity 99.7%; Pred. No. 4.2e-216;
Matches 576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 NKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKATEALS 61
DB 2 NKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSEWALKATEALS 61
QY 62 ELHGKPIEVESHVPKQRIKQIRNIPPHLQWVDSLVVQGVVSCQVNTDSETA 121
DB 62 ELHGKPIEVESHVPKQRIKQIRNIPPHLQWVDSLVVQGVVSCQVNTDSETA 121
QY 122 VNYTYSKDOARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLQOPRGRGLGQSS 181
DB 122 VNYTYSKDOARQALDKLNGFQLENFTLVKVAIPDEMAAQNPLQOPRGRGLGQSS 181
QY 182 OGSPGSVKQKPCDPLRLVPTQFVGAIIIGEGATIRNITKQTSKIDVHRKENAGAA 241
DB 182 OGSPGSVKQKPCDPLRLVPTQFVGAIIIGEGATIRNITKQTSKIDVHRKENAGAA 241
QY 242 KSTILSTPEGSTSAACKSILEIMHKEAQDIKFTTEIPLKILAHNNFVGRILGEGRNK 301
DB 242 KSTILSTPEGSTSAACKSILEIMHKEAQDIKFTTEIPLKILAHNNFVGRILGEGRNK 301
QY 302 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 361
DB 302 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 361
QY 361 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 428
DB 361 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 428
QY 428 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 481
DB 428 KIEQDQTKITISPLQELTYLNPRTITVKNVETCAKAEIEIMKKIRESYENDIASMNL 481

```

QY 482 PKEVKLEAHIRVPSAAGRVIGKGTNVLONLSSAEVVVPRDTPDENDQVVKITG 541
DB 489 PREVKLEAHIRVPSAAGRVIGKGTNVLONLSSAEVVVPRDTPDENDQVVKITG 548
QY 542 HFYACQVAKRIQELITQVKHQKQKALQSGPPQSRK 579
DB 549 HFYACQVAKRIQELITQVKHQKQKALQSGPPQSRK 586

RESULT 10
US-09-897-778-427
: Sequence 427, Application US/09897778
: Patent No. US20020147143A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Tongtong
: APPLICANT: Marnerakis, Margarita
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedrick, Thomas S.
: APPLICANT: Cartier, Darrick
: APPLICANT: Watanabe, Yoshihiro
: APPLICANT: Henderson, Robert A.
: APPLICANT: Peckham, David W.
: APPLICANT: Fanger, Neil
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.455C16
: CURRENT FILING DATE: 2001-06-28
: NUMBER OF SEQ ID NOS: 467
: SOFTWARE: aslseq for Windows Version 4.0
: SEQ ID NO: 427
: LENGTH: 586
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-897-778-427

Query Match 99.4%; Score 2938; DB 10; Length 586;
Best Local Similarity 99.7%; Pred. No. 4.2e-216;
Matches 576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFCVDCPDDESNAKATEALSGL 61
DB 9 NKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFCVDCPDDESNAKATEALSGL 68
QY 62 ELHGKPIEVEHSVPKQRIKRIQIRNIPPHLQWVLDLQVGVVSCQVNTDSETAV 121
DB 69 ELHGKPIEVEHSVPKQRIKRIQIRNIPPHLQWVLDLQVGVVSCQVNTDSETAV 128
QY 122 VNVYSSKQDQARQALDKNGFLENFTLVAYIPDEMAAQNPLQOQPRGRLGQSGSR 181
DB 129 VNVYSSKQDQARQALDKNGFLENFTLVAYIPDEMAAQNPLQOQPRGRLGQSGSR 188
QY 182 QGSPGVSKQPCDPLRLVPTQFVGAIIKEGATIRNITKQSKIDVHRKENAGAAE 241
DB 189 QGSPGVSKQPCDPLRLVPTQFVGAIIKEGATIRNITKQSKIDVHRKENAGAAE 248
QY 242 KSTILSTPECTSAACKSILEIMHKEAODIKETEEIPKLTAHNHFNVLGKEGENLAK 301
DB 249 KSTILSTPECTSAACKSILEIMHKEAODIKETEEIPKLTAHNHFNVLGKEGENLAK 308
QY 302 IEQDQDTKITITSLPQELTYNPERTITVKNVETCAKAEIMKKIRESYENDIASMNLQ 361
DB 309 IEQDQDTKITITSLPQELTYNPERTITVKNVETCAKAEIMKKIRESYENDIASMNLQ 368
QY 362 AHLIPGLNLALGLFPPTSGMPPTSGPPSAMTPPYQFQSETETVHQTIPALSYGAI 421
DB 369 AHLIPGLNLALGLFPPTSGMPPTSGPPSAMTPPYQFQSETETVHQTIPALSYGAI 428
QY 422 KGOGHRIKQSRFAGASIKIAPAEAPDAKVRVVIITGPPPAQKAGRIYKKEENFVS 481
DB 429 KGOGHRIKQSRFAGASIKIAPAEAPDAKVRVVIITGPPPAQKAGRIYKKEENFVS 488

```

```

QY 482 PKEVKLEAHIRVPSAAGRVIGKGTNVLONLSSAEVVVPRDTPDENDQVVKITG 541
DB 489 PREVKLEAHIRVPSAAGRVIGKGTNVLONLSSAEVVVPRDTPDENDQVVKITG 548
QY 542 HFYACQVAKRIQELITQVKHQKQKALQSGPPQSRK 579
DB 549 HFYACQVAKRIQELITQVKHQKQKALQSGPPQSRK 586

RESULT 11
US-09-873-637-2
: Sequence 2, Application US/09873637
: Patent No. US20020061543A1
: GENERAL INFORMATION:
: APPLICANT: Ross, Jeffrey
: TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
: FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
: CURRENT APPLICATION NUMBER: 960296.95131
: CURRENT FILING DATE: 2001-06-04
: NUMBER OF SEQ ID NOS: 46
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 577
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-873-637-2

Query Match 74.1%; Score 2190; DB 10; Length 577;
Best Local Similarity 74.1%; Pred. No. 5.1e-159;
Matches 434; Conservative 62; Mismatches 74; Indels 16; Gaps 8;

QY 1 MKKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFCVDCPDDESNAKATEALSGL 60
DB 1 MKKLYIGNLSENAAPSDLESIFKDAKIPVSGPFLVKTGYAFCVDCPDDESNAKATEALSGL 60
QY 61 IEHGKPIEVEHSVPKQRIKRIQIRNIPPHLQWVLDLQVGVVSCQVNTDSETA 120
DB 61 VELOKRLMEHSEVPKQSRKRIQIRNIPPHLQWVLDLQVGVVSCQVNTDSETA 120
QY 121 VNVYSSKQDQARQALDKNGFLENFTLVAYIPDEMAAQNPLQOQPRGR-GLGQGS 179
DB 121 VNVYSSKQDQARQALDKNGFLENFTLVAYIPDEMAAQNPLQOQPRGR-GLGQGS 176
QY 180 SROGSP--GSVSKQPCDPLRLVPTQFVGAIIKEGATIRNITKQSKIDVHRKEN 236
DB 177 PROGSEVAAGAPAKQOQVDPDPLRLVPTQFVGAIIKEGATIRNITKQSKIDVHRKEN 236
QY 237 AGAAEKSTITLSTPECTSAACKSILEIMHKEAODIKETEEIPKLTAHNHFNVLGKEG 296
DB 237 AGAAEKSTITLSTPECTSAACKSILEIMHKEAODIKETEEIPKLTAHNHFNVLGKEG 296
QY 297 RNLKXIEDQDTKITITSLPQELTYNPERTITVKNVETCAKAEIMKKIRESYENDIA 356
DB 297 RNLKXIEDQDTKITITSLPQELTYNPERTITVKNVETCAKAEIMKKIRESYENDIA 356
QY 357 SNLQAHILPGLNLALGLFPPTSGMPPTSGPPSAMTPPYQFQSETETVHQTIP 413
DB 357 SNLQAHILPGLNLALGLFPPTSGMPPTSGPPSAMTPPYQFQSETETVHQTIP 413
QY 414 ALSVGAIIKQGOHRIKQSRFAGASIKIAPAEAPDAKVRVVIITGPPPAQKAGRIYK 473
DB 414 AAVGAIIKQGOHRIKQSRFAGASIKIAPAEAPDAKVRVVIITGPPPAQKAGRIYK 473
QY 474 IREENFVSPKEEYKLEAHIRVPSAAGRVIGKGTNVLONLSSAEVVVPRDTPDEND 533
DB 474 LKEENFVSPKEEYKLEAHIRVPSAAGRVIGKGTNVLONLSSAEVVVPRDTPDEND 533
QY 534 QVVKITGHFVACQVAKRIQELITQVKHQKQKALQSGPPQSRK 579
DB 534 QVVKITGHFVACQVAKRIQELITQVKHQKQKALQSGPPQSRK 577

```

```

RESULT 12
US-09-764-864-1116
; Sequence 1116, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1116

Query Match          65.8%; Score 1946.5; DB 10; Length 620;
Best Local Similarity 65.1%; Pred. No. 2.2e-140;
Matches 385; Conservative 81; Mismatches 96; Indels 29; Gaps 8;

QY 1 MNKLYIGNLSNAAPSDLESIFKDAKIPVSGPFLVKTGYAFVDCPDSDWALKATEALSGK 60
DB 23 MNKLYIGNLSNAVPAADDRLQFGRDKPLAGVQLKSGYAFVDPDONWAIKRALETLSGK 82
QY 61 IELHGKPIEVEHSPKQRIKRLQIRNIPPHLOWELVDSLVOYGVSECEQVNTDSETA 120
DB 83 VELHGKIMEVDYSYSSKLRSRKIQIRNIPPHLOWELVDGLLQAYGTVENVEQVNTDTEA 142
QY 121 VVNTYSSKQOARALDKLNGFQLENTLVKAYIPDEMAAQNPLQPRGRGLGQRSS 180
DB 143 VVNTYATREAAKAMEKLSGHQFENYSFKISYIPDEVSPPSPQ--RAQR--GDHSSR 198
QY 181 RQG-SPGSVSKQKCDPLRLVLTQFVGALIGKEGATIRNITKQSDIDVHRKENAGA 239
DB 199 EQGHAPGTSQARQIDFLRLVLTQFVGALIGKGLTKNITKQTSRVDIHRKENSAGA 258
QY 240 AEKSTILSTEGTSAACKSILEIMHKAQDIKFTTEIPLKLAHNNFVGLIGKEGRNL 299
DB 259 AEKPTIHATPEGTSEACRMILETMQKDETKLAEEIPLKLAHNLVGLRIGKEGRNL 318
QY 300 KKIEQDTTKITISPLQELTLYNPERTITVKGNETCAKAEETMKKIRSYENDIASMN 359
DB 319 KKIEHETTKITISLQDLSYNPERTITVKGTEACASAEIEMKKLREAFENDMLAVN 378
QY 360 LQAHILPGLNLNALGLF-----PPTS--GMPPT-----SGPPSAMTP-----P 396
DB 379 QOANILPGLNLSALGIFSTGLSVLSPAGPAGPAPYHPTTHSGYFSSLYPHHFGP 438
QY 397 YPQFEQ-SETETVHOFTPALSVGAIIGKQGHKOLSRFAGASIKIAPAEAPDAKVRMI 455
DB 439 FPHHSYPEQEIVNLFITQAVGAIIGKKGAGHAKOLARFAGASIKIAPAEQPDVSRMI 498
QY 456 ITGPPPEAQFKAQRIYGIKEENFVSPKEEVKLEAHIRVPSFAGRVIGKGGKTVNELQN 515
DB 499 ITGPPPEAQFKAQRIYGIKEENFVSPKEEVKLEAHIRVPSFAGRVIGKGGKTVNELQN 558
QY 516 LSSAEVVPDQTPDNDQVVKITGHFYACQVAQRKIQEITLTQVKHQHQ 566
DB 559 LTSAEVVPDQTPDNDQVVKITGHFYACQVAQRKIQEITLTQVKHQHQ 609

RESULT 13
US-09-764-864-1114
; Sequence 1114, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1114
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1114

Query Match          31.1%; Score 919; DB 10; Length 250;
Best Local Similarity 73.3%; Pred. No. 1.6e-62;
Matches 187; Conservative 26; Mismatches 34; Indels 8; Gaps 5;

QY 328 TVKGNVETCAKAEETMKKIRSYENDIASMNLAHLIPGLNLNALGLFPPTSGMPPTS 387
DB 1 TVKGATENCRAEQEIMKKVREAYENDVAMSLQSHLIPGLNLAAVGLFPASSAVPP-- 58

```

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1114
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1114

Query Match          32.1%; Score 949; DB 10; Length 261;
Best Local Similarity 74.5%; Pred. No. 8.5e-65;
Matches 193; Conservative 26; Mismatches 32; Indels 8; Gaps 5;

QY 324 ERTITVKGNETCAKAEETMKKIRSYENDIASMNLAHLIPGLNLNALGLFPPTSGMP 383
DB 8 ERTITVKGATENCRAEQEIMKKVREAYENDVAMSLQSHLIPGLNLAAVGLFPASSAV 67
QY 384 PPTSGPPSAMT--PPYPQFEQS-ETETVHOFTPALSVGAIIGKQGHKQLSRFAGASIK 440
DB 68 PP---PPSYTGAAVYSSSPMQAPEQEMVQVFPQAQVGAIIIGKQGHKQLSRFASASIK 124
QY 441 IAPAEAPDAKVRMVIITGPPEAQFKAQRIYGIKEENFVSPKEEVKLEAHIRVPSFAAG 500
DB 125 IAPETPDSKVRMVIITGPPEAQFKAQRIYGIKEENFVSPKEEVKLEAHIRVPSAAG 184
QY 501 RVITGKGKTVNELQNLSAEVVPDQTPDNDQVVKITGHFYACQVAQRKIQEITLTQV 560
DB 185 RVITGKGKTVNELQNLSAEVVPDQTPDNDQVVKITGHFYACQVAQRKIRDILOV 244
QY 561 KOHQOQKALQSGPPQSRKK 579
DB 245 KQ-QHQGQ-QSNOQAQARRK 261

RESULT 14
US-09-764-864-1532
; Sequence 1532, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1532
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1532

Query Match          31.1%; Score 919; DB 10; Length 250;
Best Local Similarity 73.3%; Pred. No. 1.6e-62;
Matches 187; Conservative 26; Mismatches 34; Indels 8; Gaps 5;

QY 328 TVKGNVETCAKAEETMKKIRSYENDIASMNLAHLIPGLNLNALGLFPPTSGMPPTS 387
DB 1 TVKGATENCRAEQEIMKKVREAYENDVAMSLQSHLIPGLNLAAVGLFPASSAVPP-- 58

```


Qy 411 ---FIPALSVGAIGKQGHQHIKOLSRFAGASTKIAPAEAPDAKV----- 451
 Db 569 VRMVPSMIGALIGAKGNKIKIIRDTGASVKI---EAPKEKTQREAEAEKRRKLDST 625
 Qy 452 -----RWLIITGPPEAKQAGRIYXK 473
 Db 626 DSGCEGVASGDHPQEFLEDNATINSSDAIEEKPVSERWVINGDDQLLQAQYVFSK 685
 Qy 474 IKEENFYSP-----KEEVKLEAHIRVPSFAAGRVIGKGTNVLONLSSAEVVPVR 525
 Db 686 IAEITSSSLPSSGMDGRSHMLRITRTEVSPTRIIGRIIGKGNVRELQRTGAVVVKIPE 745
 Qy 526 DQTPD-----ENDQVVVKITGFYACQVAQRKIQEILTOVKQ---HQQAKA-LQS 571
 Db 746 EERNNGEVYRHDDGLEEDMTNITIGNMYSTHNVPQRLAHLVNEYRSGDHNRKSSDYK 805
 Qy 572 GPPOS 576
 Db 806 GRPHS 810

RESULT 2
 T27855
 hypothetical protein ZK418.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27855
 R:Fulton, L.
 A:Description: The sequence of C. elegans cosmid ZK418.
 A:Reference number: Z20430
 A:Accession: T27855
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-768 <FULL>
 A:Cross-references: EMBL:U00047; PIDN:AAA50693.1; CESP:ZK418.9
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:ZK418.9
 A:Introns: 166/1; 231/3; 243/3; 279/1; 387/3; 490/1; 567/3; 727/2

Query Match 8.2%; Score 243.5; DB 2; Length 768;
 Best Local Similarity 21.2%; Pred. No. 9.2e-08;
 Matches 123; Conservative 103; Mismatches 215; Indels 139; Gaps 22;

Qy 62 ELHGKPIEVSVPKQRIKRLQIRNIPPHLQWEVLDSLLVQGVVSECEVNTDSEAV 121
 Db 139 QLIKGLIELD--ITKNESL-----LLSALFL-----CSAVHPDIDSTEG 174
 Qy 122 VNVITYSSKDQARQAL-----DKLNGFQ-LENFTLKVAYIPDEMAAQON 163
 Db 175 KNLKYQYQYSAALINHCCLTNQONAPSRYSSELLSYQIIEETHQKIIGADGGIGTLKR 234
 Qy 164 PLQQ-----PGRRLGQSSSGSGSVSKQKPCDPLRLLLVPTQFVGAIGREG 215
 Db 235 PLDSEILDGLDLPFKKS--SEVGDNLNGDSKDTIDYIP-----VPEKVVGLVIGKGG 284
 Qy 216 ATIRNITKQTSKIDVHRKENAGAAKSITILSTPGTSAAKSIILEIMHKE---AQDIK 272
 Db 285 SEIRLIQOTSCRVQMPDPHQSVNGFNCTIEGPPQVAVAROMITQVLRNQTGAQPCA 344
 Qy 273 FTEEIPKILAHNVFGRLLGKGRNLIKKEQDQTDKTIISPLQELTYL-----NPERTIT 328
 Db 345 APGEVTEMLTPADKIGLVIGKGETIRIVQESGLR-NCNVVQETTTATGQPKPLRMIG 403
 Qy 329 VKNVETCAKAEIEIMKRESYENDIASNNLQAHILPGLNLAALGFLPPTSGMPPPTSG 388
 Db 404 SPAAIET-AXA-----LVHNMNNTQGNAPLQORAPHQPSG 438
 Qy 389 PPSAMTPYPQF-----EQSETVTHQFIPALSVGAIGKQGHQHIKOLSRFAGASTK 440
 Db 439 -----QFGGGAQEAQAKGEVI---VPRLSNGLMIGKGMIEKRLAAETGTKIQ 485

Qy 441 IAPAEAPDAKVRMVIITGPPEAQFKAQRIYKIKE---ENFVSPKEEVKLEA----- 490
 Db 486 FKPDTPNPSEDRIATVMTGRDQIYRATERITEIVNRAIKNNGAPQDRGSACTVLPQOSIF 545
 Qy 491 HIRVPSFAAGRVIGKGTNVLONLSSAEV-VVPRDQTPDENQDVVKITGH-----FYA 545
 Db 546 YMHVPACKCGLVIGKGGENIKQIERETGATCGLAPEAEQKNE-DEKVFEEKSLOLIHHA 604
 Qy 546 CQVAQRKIOBILFQVK-----QHQQAKALQSGPPQS 576
 Db 605 SHLVRIKVGESIPNTVPVPLQAGGGYQOQQAAMFASGTON 644

RESULT 3
 A53184
 myc far upstream element-binding protein - human
 N:Alternate names: FUSE-binding protein
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
 C:Accession: A53184
 R:Duncan, R.; Bazar, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan, M.; Leve
 Genes Dev 8; 465-480, 1994
 A:Title: A sequence-specific, single-strand binding protein activates the far upstream
 A:Reference number: A53184; MUID:94170991; PMID:8125259
 A:Accession: A53184
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-644 <DUN>
 A:Cross-references: GB:U05040; NID:g460151; PIDN:AAAL1976.1; PID:g460152
 C:Keywords: DNA binding

Query Match 8.1%; Score 240; DB 2; Length 644;
 Best Local Similarity 22.3%; Pred. No. 1.2e-07;
 Matches 100; Conservative 75; Mismatches 154; Indels 120; Gaps 17;

Qy 161 QONPLQ---OPRRGRLGQSGSRQSGSVSKQKPCDPLRLLLVPTQFVGAIGKEGAT 217
 Db 63 OKRPLEDGDQDAKKVAPQNDSPGTQLPMHQOQSSVMTEEVKVPDGMVGFIIGRGGEQ 122
 Qy 218 IRNITKQTSKIDVHRKENAGAAKSITILSTPGTSAAKSIILEIMHKEAQDIKFE-- 275
 Db 123 ISRIQESGCKIQI-APDSGGLPERSCLMTGTTPESVQSAKRLDQIVKGRPAQGFHHGD 181
 Qy 276 ---EIPKILAHNVFGRLLGKGRNLIKKEQDQTDKTIISPLQELTYLNPRT-----I 327
 Db 182 GPGNAVQEIIMPASKAGLVIGKGETIKQLQERAGVKVM--IQD---GPONTGADKPL 235
 Qy 328 TVKGNVETCAKAEIEIMKIRE-----SYENDIASNNLQAHILPGLNLAALGFLPPTSGM 382
 Db 236 RITGDPYKVOQAKEMVLELIRDQGGFREVNEYGSR-----IGG-----NEGI 278
 Qy 383 PPPTSGPPSAMTPYPQFQSEETVHQFIPALSVGAIGKQGHQHIKOLSRFAGASIKIA 442
 Db 279 DVP-----IPRFAVGIVIGRNGEMIKKIQNDAGVRIQFK 312
 Qy 443 PAEAPDAKVRMVIITGPPE-AQPKAQ-----GRIYK----- 473
 Db 313 PDDGTTPE-RIAQITGPPDRCOHAAEITDLLRSVQAGNPGPGGGRGRGRCQGNWNG 371
 Qy 474 ---IKEENFVSPKEEVKLEAHIRVPSFAAGRVIGKGTNVLONLSSAEVVPVDDQTP 529
 Db 372 PPGGLQFENFI-----VPTGKTLIGKGETIISQSQSGARIELORNPFP 418
 Qy 530 DENDQV-VVAKITGH---FYACQVAQRKI 553
 Db 419 NADPNMKLFTIRGTPOQIDYAROLIEKI 447

RESULT 4
 T49962
 hypothetical protein F9M21.160 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49962
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 224493
 A:Accession: T49962
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-568 <BEV>
 A:Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160
 A:Experimental source: cultivar Columbia; BAC clone F8M21
 C:Genetics:
 A:Gene: ATSP:F8M21.160
 A:Map position: 5
 A:Introns: 200/3; 337/3; 544/3

Query Match 8.0%; Score 236.5; DB 2; Length 568;
 Best Local Similarity 22.8%; Pred. No. 1.7e-07;
 Matches 92; Conservative 79; Mismatches 144; Indels 89; Gaps 16:

QY 195 DLPLRLVPTQFVGAIGKEGATIRNITKQTSKIDVHRKENAGAAKSTITLSTPEGTS 254
 Db 49 DTVRYLCPVKKIGSVIGRGDIVKOLNDRTRSKIRI-GEALPGCDERVIITVSPSDET 107
 QY 255 A-----ACSIIEIMHKAQDIKFEIP-----LKIANNFVGLIGKEGR 297
 Db 108 AFGDEKVLSPAQDALFRIDRVVADDAKSEDSPEGEKVQAKLLVSDQIGCLGRGGQ 167
 QY 298 NLKKIEQDTRKTIISPLQELT--YNPERTIVKGNVETCAKAEEMKKIRESYENDI 355
 Db 168 IVONIRSETGAQIRIVDRNMPLCALNSDELQISGEVLIVKALLQIASRL---HENPS 224
 QY 356 IASNNLOAHILPGLNLAGLFPPTS-----GMPP-----PTSGPP 390
 Db 225 RSONL-----LSSGGYPAGSLMSHAGGPRVLGLAPLMGVSGRDAGDWSRPLYQPP 275
 QY 391 SAMTPPYQFQOSTETVHOFIPALSVGLIGKQGHKOLSRPAGASIKI--APAPAD 448
 Db 276 -RNDPPATF-----FIRLVSVENIASVIGKGGALINQLRQETRAIKVDDSRTEGND 328
 QY 449 AKVRVVIITGPP-----EAFQKAGRIYGIKKEENFVSPKEEVKLEAHIRVPSFA 498
 Db 329 C---LITTSAREVEFEDAYSPTIEAVMLQPKCSKVERDSGL-----VSFTRLLVPSR 380
 QY 495 AGRVIGKGGKTVNELONLSSAEV--VPRQTP-----DENDQVVV 537
 Db 381 IGCILGKGGAITEMRRTKANTIRILGKENLPKVASDDDEMVOV 424

RESULT 5
 D96554
 Hypothetical protein F19C24.19 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: D96554
 R:theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D96554
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-621 <SNO>
 A:Cross-references: GB:AE005173; NID:g11094762; PIDN:AAG29695.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F19C24.19
 A:Map position: 1

Query Match 8.0%; Score 235.5; DB 2; Length 621;
 Best Local Similarity 22.9%; Pred. No. 2.2e-07;
 Matches 104; Conservative 76; Mismatches 171; Indels 103; Gaps 17:

QY 199 RLLVPTQFVGAIGKEGATIRNITKQTSKI-----D 230
 Db 22 RLLCPATRTGAIIGKGGVIRHLQSVTGSKIRVIDIPVSEERVVLIAPSKKKDESN 81
 QY 231 VHRKENAGA---AEKSITILST---PEGTSAACKSILEIMHK-----EAQDI----- 271
 Db 82 VCDSENPGSEEPKQKSGECAGTSGGDDEAPSSAQMALLRVFERIVFGDDAATVGDDEL 141
 QY 272 -KFTEEIDPLKILAHNNFVGLIGKEGRNLKKIEQDTRKTIISPLQEL--TLNPERTIT 328
 Db 142 DKGESEGLRMVGNQVDYIMSKRGKMKIKRSDGAIVRISDQIPPCAPFGDVVQ 201
 QY 329 VKGNETCAKAEEMKKIRES-----YENDIASNNLOAHILPGLNUNAL 373
 Db 202 MNGKFSVKKALLVTLNCLQESGAPPTWDECPFPQPGPYEYHSM-----247
 QY 374 GLFPPTSGMPPPTSGPPSAMTPPYQFQOSTETVHOFIPALSVGLIGKQGHKOLSR 433
 Db 248 --YHPQWDHPPPNMPED--VGPPNRPVVEEVAERLLCPADKVGSLIGKGAIVRAALQN 303
 QY 434 FAGASIKIAPAEADPAKRVVIITGPEAQK---AQG---RIYKIKEENFYSPKEEVK 487
 Db 304 ESGASIKVSD--PTHDSERIIVISARENLERHSLAOGVMVHNRIVETGF---EPSAA 359
 QY 488 LGAHVRVPSFAAGRVIGKGGKTVNELONLSSAEV--VPRQTPDENDQ--VVVKITGHFY 544
 Db 360 VVARLLVSHPVIGRLGKGGLHISEMRATCASIRVFAKDQATKYESQHDREIVQVIGNL- 418
 QY 545 ACQVQRKIQLTQVQKQKQKAL---QSGPP 574
 Db 419 --KTQDVALFQLCLRLEAMFPGLPQMGSGPP 450

RESULT 6
 T48439
 probable RNA-binding protein - Arabidopsis thaliana
 N:Alternate names: protein T32M21.30
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48439
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.;
 ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: 224467
 A:Accession: T48439
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-313 <BEV>
 A:Cross-references: EMBL:AL162875
 A:Experimental source: cultivar Columbia; BAC clone T32M21
 C:Genetics:
 A:Map position: 5
 A:Introns: 29/1; 108/3; 144/2; 209/1; 229/3; 261/3; 282/2
 A:Note: T32M21.30

Query Match 7.8%; Score 231.5; DB 2; Length 313;
 Best Local Similarity 25.1%; Pred. No. 1.4e-07;
 Matches 79; Conservative 62; Mismatches 131; Indels 43; Gaps 10:

QY 183 GSPGVSQKQP-----CDLP--LRLVPTQFVGAIGKEGATIRNITKQTSKID 230
 Db 10 GSPEELAKRSPHDSSEADSAEKPTIRFLVSNAAAGSVIGKGGSTITEFOAKSGARIQ 69
 QY 231 VHRKEN--AGAAEKSTITLSTPEGTSAACKSILEIMHKAQDIKFEIP---LKIHAH 285
 Db 70 LSRNQEFFGTDRILMISGSIKEVGVNGLLEILDKLHSELHAEDGNEVEPRRLRLVFN 129
 QY 286 NFVRLIGKEGRNLKKIEQDTRKTIISPLQELTLN--PERTITVKGNETCAKAEEM 344

Db 130 SSCGIIKGATATKSFTEESKAGIKISPLDN-TFYGLSDRLVLTSGTFEOMRAIDLIL 188
 Qy 345 KKIRESYENDIASMNIQAHLIPGLNALGLFPTSGMPPTSGPPSAMTPPYFOFQSE 404
 Db 189 AKLT--EDDHYSQNVHSPSYAGYNSVNPAPNGS-----GKYQNHKEEA 232
 Qy 405 TETVHOFIPALSVGATIGQGHQKLSRFAGASIKIAP-----AEAPDAKVRVITGP 459
 Db 233 STVTIGVADERHIGLVGRGNIMEITQWTGARIKISORGDMSCTTDRKVS---ITGP 289
 Qy 460 PEQFKAQGIYKI 474
 Db 290 QRAIQAEHWIKOV 304

RESULT 7
 T41600
 probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 01-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T41600
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z22003
 A:Accession: T41600
 A:Status: preliminary; translated from GB/EMBL/DBSJ
 A:Molecule type: DNA
 A:Residues: 1-398 <LYN>
 A:Cross-references: EMBL:AL01825; PIDN:CAA21234.1; GSPDB:GN00068; SPDB:SPC757.09c
 A:Experimental source: strain 972h-; cosmid c757
 C:Genetics:
 A:Gene: SPDB:SPC757.09c
 A:Map position: 3

Query Match 7.8%; Score 229.5; DB 2; Length 398;
 Best Local Similarity 26.5%; Pred. No. 2.8e-07;
 Matches 90; Conservative 55; Mismatches 148; Indels 47; Gaps 11;
 Qy 176 GRSSRGSGSGS-----VSKQKPCD-----LPLRLVPTQVGAIGKEGATIRNITK 223
 Db 62 POSMTNPPTSVOVPPISAKPPMDATYATQQLTLRALSTREAGIIGKAGKNVAELRS 121
 Qy 224 QTSKIDVHRKENAGAEEKSTITLSTPEGTSAAKSTILEIMHKEADIKFTEE-----IP 278
 Db 122 TTNVKAQV-TKAVPNVHDRVLTISGLENVVRAVRFIDIFAKNSTNPDGTPSDANTPRK 180
 Qy 279 LKILAHNNFVGRIGKEGNLKKIEODTDTKITISPLQELTYNPERTITVKGNETCAK 338
 Db 181 LRLIASHLSGSIIGRNLRIKIDKSCRMAS--KDLPOSTERTVEIHGTVDNLHA 238
 Qy 339 ABEEMKKIRESYENDIASM--NLQHL---IPGLNLNA--LGLFPTSGMPPTSGP-- 389
 Db 239 AIWEIGKCLIDDERGAGTVFVNPVSRITQPLSLASTASPOQVSPAA--PSTTSGEAI 296
 Qy 390-----PSAMTPPYPOFQSESETVHOFIPALSVGATIGQGHQKLSRFAGA 437
 Db 297 PENFVSYGAQVFATOMP-----FLQPKVTQNTISIPADMVGCILIGSGSKISEIRTSGS 352
 Qy 438 SIKIAPAEAPDAKVRVITGPPEAOKFAQGIYKIKE 477
 Db 353 KISIAKEPHDTCERMTITGTGHEENKALEFLYQOLEME 392

RESULT 8
 138489
 onconeural ventral antigen-1 - human
 C:Species: Homo sapiens (man)
 C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
 C:Accession: 138489
 R:Buckanovich, R.J.; Posner, J.B.; Darnell, R.B.
 Neuron 11, 657-672, 1993
 A:Title: Nova, the Paraneoplastic R1 Antigen, is Homologous to an RNA-Binding Protein and

A:Reference number: 138489; MUID:94000830; PMID:8398153
 A:Accession: 138489
 A:Status: preliminary; translated from GB/EMBL/DBSJ
 A:Molecule type: mRNA
 A:Residues: 1-510 <RES>
 A:Cross-references: EMBL:U04840; NID:g440877; PIDN:AAA16022.1; PID:g440878
 C:Genetics:
 A:Gene: Nova-1

Query Match 7.7%; Score 227; DB 2; Length 510;
 Best Local Similarity 22.4%; Pred. No. 5.8e-07;
 Matches 119; Conservative 77; Mismatches 179; Indels 156; Gaps 20;
 Qy 155 PDMAAQNPLOQPRGRGLGQRRSGSGSPGVSQKPCDPLPLLLVPTQFVGAIGKE 214
 Db 23 PD---SRKRPLEAP-----PEAGSTKRTNTGEGQYF-----LKVLPSTYAGSIIGKG 68
 Qy 215 GATIRNITKQTSKIDVHRKENA-----GAAEKSTITLSTPEGTSAAKSTILEIMHKEAQ 269
 Db 69 GQIVQLQKETGATIKLSKSKDFYPTGTVTRVCLIOGTVEALNAVHGFTAEREMPO 128
 Qy 270 DIKFTEEIPL-----KILAHNNFVGR 290
 Db 129 NVAKTEPVSILOPQTVNPDRIKOTLPSSPTTTKSSPSDPMTTSRANQVKIIVPNSTAGL 188
 Qy 291 LIKCEGRNLKKIEODTDTKITISPLQELTYNPERTITVKGNETCAKABEIMKKIRE- 349
 Db 189 IIRKGGATVAVNEOSGAWVOLSQKPD-GINLQERVTVVSGEPQNRKRAVELIIQIKQED 247
 Qy 350 -----SYEN--DIASMN-----LQAHLPGLNLNALGLFPT- 379
 Db 248 POSGSCNLISYANVTGPVANSNPTSGPYANTAEVLTAAAGLGHANLAGVAAPFAVL 307
 Qy 380 SGMPPPTSGPPSAMTPPYPOFQSESETVHOFIPALSVGATIGQGHQKLSRFAGASI 439
 Db 308 SGP---TGNDLVAIT-----SALNTL-----ASYCYNLTNLGLSLQAAA-TGALA 349
 Qy 440 KIAPAEAPDAKVRVITGPPEAOKFAQGIYKIKI----- 474
 Db 350 AAAASANPAAAAANLATAIYASEA--SASGSTAGTAGTATGALGSLAAATAATNGYFGAASP 407
 Qy 475 -----KEENFVSPEEVEKLEAHIRVPSFAAGRVIGKGTGKTVNELQNLSABVVVPR- 525
 Db 408 LAASATLGTERTKSDGSKDVE---IAPENLVGAILGKGTGLVEYQELTGARIQISKK 463
 Qy 526 -DQTPDENQVVKITIGTHFYACQVAKRQIEILTVQKQHQOQKALQSGPPQ 575
 Db 464 GEFVPGTRNRKVT-ITGTPAATAQAQYLITORTIT-----YEQGVRAANPQ 507

RESULT 9
 T19216
 hypothetical protein cl2D8.1a - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19216
 R:McMurray, A.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: Z19092
 A:Accession: T19216
 A:Status: preliminary; translated from GB/EMBL/DBSJ
 A:Molecule type: DNA
 A:Residues: 1-589 <WIL>
 A:Cross-references: EMBL:Z73969; PIDN:CAA98232.1; GSPDB:GN00023; CESP:CL2D8.1a
 A:Experimental source: clone C12D8
 C:Genetics:
 A:Gene: CESP:CL2D8.1a
 A:Map position: 5
 A:Introns: 7/3; 25/1; 159/3; 318/3; 513/2

Query Match 7.6%; Score 225.5; DB 2; Length 589;
 Best Local Similarity 23.4%; Pred. No. 8.9e-07;
 Matches 96; Conservative 72; Mismatches 157; Indels 53; Gaps 11;

QY 202 VPTQFVGAIGKAGATIRNITKQTSKIDVHRKENAGAAEKSTILSTPEGTSAAKSKIL 261
 DB 54 IPESAVGIVIGGSGEIOGQAKACRVOMSPDADSSGVWVMTLGRSNVETAKHLIN 113
 QY 262 EIMHKAODIKFTEIP-----LKILAHNFVGRLLIGKEGRNLKKIEQDQDTK-ITISPL 315
 DB 114 EVVAR-SONPRPQYGFPRATTDIAIPNRCGLIIGKSGDTIRLOQKSGCKMILVQDN 172
 QY 316 QELTYLNPERITVYKGNVETCAKAEIEIMKKIRSYENDIASMNLQAHLPGLNLNALGL 375
 DB 173 QSVS--DOSPLRITGDPQKIELA-KQLVAEILNSGGDNGSGSLQMHAGG----- 221
 QY 376 FPPTSGMPPTSGPPSAMTPPYQFEQSETETVHOFPALSVGAITKQGOHIKQLSRFA 435
 DB 224 -----GGGASA-----RGEVV--VPRSSVGIIGKQGDITIKRLAMET 256
 QY 436 GASIKIAPAEAPDAKVRMVIITGPPPEAOFKAGRIYKIKIENFVS-----PKEEVKLE 489
 DB 257 GTKIOFKPDDPSTPERCAVIMGTROQIVRATERITELVKKSTMOQGGGNGVAGAMVSN 316
 QY 490 A---HIRVPSFAAGRVIGKGVNVLONLSSAEVVPDQTPDENDOVVVKITGHFYAC 546
 DB 317 ASTYMSVPAKAGLVIGGGETIKQINSSEGAHCELSRDTGNADEKVFV-IKGKRAI 375
 QY 547 QVAQRKIO 554
 DB 376 EHAKHLIR 383

RESULT 10
 T19217
 hypothetical protein C12D8.1b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 R:McMurray, A.
 A:Submitted to the EMBL Data Library, June 1996
 A:Reference number: Z19092
 A:Accession: T19217
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-611 <WTL>
 A:Cross-references: EMBL:Z73969; PIDN:CAA98233.1; GSPDB:GNO0023; CESP:C12D8.1b
 A:Experimental source: clone C12D8
 C:Genetics:
 A:Gene: CESP:C12D8.1b
 A:Map position: 5
 A:Introns: 181/3; 340/3; 535/2

Query Match 7.6%; Score 225.5; DB 2; Length 611;
 Best Local Similarity 23.4%; Pred. No. 9.4e-07;
 Matches 86; Conservative 72; Mismatches 157; Indels 53; Gaps 11;

QY 202 VPTQFVGAIGKAGATIRNITKQTSKIDVHRKENAGAAEKSTILSTPEGTSAAKSKIL 261
 DB 76 IPESAVGIVIGGSGEIOGQAKACRVOMSPDADSSGVWVMTLGRSNVETAKHLIN 135
 QY 262 EIMHKAODIKFTEIP-----LKILAHNFVGRLLIGKEGRNLKKIEQDQDTK-ITISPL 315
 DB 136 EVVAR-SONPRPQYGFPRATTDIAIPNRCGLIIGKSGDTIRLOQKSGCKMILVQDN 194
 QY 316 QELTYLNPERITVYKGNVETCAKAEIEIMKKIRSYENDIASMNLQAHLPGLNLNALGL 375
 DB 195 QSVS--DOSKPLRITGDPQKIELA-KQLVAEILNSGGDNGSGSLQMHAGG----- 243
 QY 376 FPPTSGMPPTSGPPSAMTPPYQFEQSETETVHOFPALSVGAITKQGOHIKQLSRFA 435
 DB 244 -----GGGASA-----RGEVV--VPRSSVGIIGKQGDITIKRLAMET 278
 QY 436 GASIKIAPAEAPDAKVRMVIITGPPPEAOFKAGRIYKIKIENFVS-----PKEEVKLE 489
 DB 279 GTKIOFKPDDPSTPERCAVIMGTROQIVRATERITELVKKSTMOQGGGNGVAGAMVSN 338

QY 490 A---HIRVPSFAAGRVIGKGVNVLONLSSAEVVPDQTPDENDOVVVKITGHFYAC 546
 DB 339 ASTYMSVPAKAGLVIGGGETIKQINSSEGAHCELSRDTGNADEKVFV-IKGKRAI 397
 QY 547 QVAQRKIO 554
 DB 398 EHAKHLIR 405

RESULT 11
 C96275
 hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, J.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo, A.; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.C.; Davis, R.W.; Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the Plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96275
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-479 <STO>
 A:Cross-references: GB:AE005172; NID:g5080792; PIDN:AAD39302.1; GSPDB:GNO0141
 C:Genetics:
 A:Map position: 1

Query Match 7.5%; Score 222; DB 2; Length 479;
 Best Local Similarity 21.8%; Pred. No. 1.1e-06;
 Matches 98; Conservative 93; Mismatches 169; Indels 94; Gaps 19;

QY 165 LOQPRRRG-----LQGRSSRQSGSPGVSKQKPC--DLPLRLVPTQFVGAILCKEGA 216
 DB 5 LRNIHGRSNLQSEFTGNGSKSRNLHDETDQNVJASEDTVRYLCPVKTKGSIIGKGE 64
 QY 217 TIRNITKQTSKIDVHRKENAGAAEKSTILSTPEGTS-----AAKSIIEIMHK--- 266
 DB 65 IAKIRSETKSNMRIN-EALPGCEERVMTMYSTNEELNHFHGDGELVCPA-LOALEKVDH 122
 QY 267 -----EAQDIKETEELPKILAHNFVGRLLIGKEGRNLKKIEQDQDTKITI- 312
 DB 123 MVVADADQDDGTDNDLGEKQIVTVRMLVPSDQIGCVIGKGGVIONLRDINAQIRVI 182
 QY 313 ---SPLQELTYLNPERITVYKGNVETCAKAEIEIMKKIRFES---YENDIASMNLQAHLP 366
 DB 183 KDHLFACALITLSDLELLII-GEPLVVRREALQVASLLHDNPSRFQHLSSSSSSSSMHOP 241
 QY 367 GLNLNALGLFPPTSGMPPTSGPPSAMTPPYQFE-QSETETVHOQ-----IPALSVGAI 420
 DB 242 GAMLMS-----AALTSSHRNYAVRRDIADAREFCVCFCIPAEVNGV 283
 QY 421 IGKQGHKQLSRFAGASIKITAPAEAPDAKVRMVIITGPE-----AOFKAGRI 470
 DB 284 IGKGGFNIQIETGATIRVNTSETDDD--CIIFISKEFEYEQSPAVNAIRLOQR 341
 QY 471 YGKI-KEENFVSPKBEVKLEAHIRVPSFAAGRVIGKGVNVLONLSSAEV-VPRDQT 528
 DB 342 SEKVQKDN-----DLAISTRLVSSSQIGCLIGKGVAVSEMSKSVTRNIRILOEDV 395
 QY 529 PD--ENDOVVVKITGHFYACQVAKIQEILTOV 560
 DB 396 PKIAREDEEMVOITGSPDAMKA-----LTQV 422

RESULT 12

S41224
hRNP protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C:Accession: S41224
R:Strom, H.; Matunis, M.J.; Michael, W.M.; Dreyfuss, G.
Nucleic Acids Res. 21, 1193-1198, 1993
A:Title: The pre-mRNA binding K protein contains a novel evolutionarily conserved motif.
A:Reference number: S41224; MUID:93219080; PMID:8464704
A:Accession: S41224
A:Molecule type: mRNA
A:Residues: 1-396 <SIO>
C:Superfamily: transformation upregulated nuclear protein HNRPK

Query Match 7.48; Score 219.5; DB 2; Length 396;
Best Local Similarity 21.3%; Pred. No. 1.2e-06;
Matches 93; Conservative 76; Mismatches 166; Indels 101; Gaps 18;

QY 175 GQSSRGSGSGVSKQKPCDL--PLRLVPTQFVGAIIGKAGATIRNITQTQSKIDVH 232
DB 17 GKRPADMEDEQAFKRSNTDMVLELRLQSKNAGAVTGGKGNKIKALRTDYNASV- 75
QY 233 RKNAGAEEKSITILSTPEGTSAAKSIILEMHKEAQDIKFTFEEPLKILAHNNFVGRLI 292
DB 76 --PDSSGPERILSISADITETIGELKLIPTLSEHFGNDF--DCELRLLIHQSLAGII 131
QY 293 GKGRNLKIEODTQKITISPLQELLYNPRTIVKGNVETCAKAEEMKIRRES-- 350
DB 132 GVAGAKIKELREKTOT--TIKLFQECPCPSTDRVLLIGGRVPRVVECIKVIDLISEPV 189
QY 351 -----YEN-DIASMNL-----QAHILPGLNINLGLF-----PPTSGMPPTS 387
DB 190 KGRSQPDYDNFYDYDYGFTWMDRRGRPH--GFSMHARGGDRMPGPGGRMPQS 246
QY 388 -----GPPSANTPPYQFQESQETVTHQFIPALSVGAIIGKQGHKQKLSRFAGAS-- 438
DB 247 RRDYDMSPRGCLPPLPP-----GRGR-----GSGRA 274
QY 439 --IKIAPAPAPAKVRNVIITGPEAQFKAQRIYKIEENFVSPKEVLEAHIRVPS 496
DB 275 RNPLPLPPPPGGGR-----RGRPDHYDGMGGRGYGRSGFDIGGP-----VITQVTIPK 326
QY 497 FAAGRVIGKGTGNELQNLSSAEVVPVPRDQTPDENDQVVKITGHYFACVQAQRKIQE- 555
DB 327 DLASGIIGKGGRIKQIRHESGASIKI--DEPLEGSDRIITITG-----TODQIONA 377
QY 556 ---ILTVQVKHQOQKA 568
DB 378 RFLQNSVKQFSEDA 393

RESULT 13
hypothetical protein M01A10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25832
R:Scheet, P.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid M01A10.
A:Reference number: Z20094
A:Accession: T25832
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-680 <SCH>
A:Cross-references: EMBL:U08174; PIDN:AA842272.1; GSPDB:GN00019; CESP:M01A10.1
A:Experimental source: strain Bristol N2; clone M01A10
C:Genetics:
A:Gene: CESP:M01A10.1
A:Map position: 1
A:Introns: 63/2; 97/3; 121/1; 160/1; 269/3; 411/3; 512/1; 649/2

Query Match 7.28; Score 213.5; DB 2; Length 680;

Best Local Similarity 21.3%; Pred. No. 6.5e-06;
Matches 73; Conservative 69; Mismatches 139; Indels 61; Gaps 9;

QY 202 VPTQFVGAIIGKAGATIRNITQTQSKIDVHRKKNAGAAKSITILSTPEGTSAAKCSIL 261
DB 152 IPESVGLVIGRNGVEIQAIQSQKSGCRVQI-VAEPSTTGTGYSVDIYGISIEVAKKLIN 210
QY 262 EIM--HKEAQQ-----IKFTFEEPLKILAHNNFVGRILGKGRNLKIEQ 304
DB 211 EVVARGKLSOEPLPCSVPOQPIPAVSNSSKVTIIPIPANKCGAIIIGKKGEMRKLR 270
QY 305 DTDTKITTSLOELTLYNPRTIVKGNVETCAKAEEMKIRSEYENDIASNLOAHL 364
DB 271 WNCDFIL--IOENNIADSVKRPQITQCPKEVEHAKA-----LVADI 310
QY 365 IFGLNINLALGLFPPTSGMPPPTGPPSAMPPTYPQFQESQETETVHOFIPALSVGAIIGKQ 424
DB 311 LQCFD-----ECPPAGMAGNSPVAAMS-----LQVKVPRSTVGAIMGLO 349
QY 425 GQHIKQLSRFAGASIKIAPAEADPAKVRNVIITGPEAQFKAQRIYKIEENFVSPKE 484
DB 350 GSNIKKISNETETKIQEPDQDDPKLMERTLVIGNKNKVVYC-ARLLQKIVEANSENANT 408
QY 485 EYKLEAHIRVSPFAGRVIGKGTGNELQNLSSAEVVPVPRD 526
DB 409 PISL-FYMLIPASKGLVIGRGETIRQINKESGAYCEMSRD 449

RESULT 14
S58529
Alpha-complex protein 1 - human
N:Alternate names: nucleic acid-binding protein; protein PCBP-1
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 01-Dec-2000
C:Accession: S58529; S58523; S58578; S43489; S41378; S42472
R:Kiladjian, M.; Wang, X.; Liebhaber, S.A.
EMBO J. 14, 4357-4364, 1995
A:Title: Identification of two KH domain proteins in the alpha-globin mRNA stability
A:Reference number: S58523; MUID:96016208; PMID:7556077
A:Accession: S58529
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-356 <KIL>
A:Accession: S58523
A>Status: preliminary
A:Molecule type: protein
A:Residues: 125-139; 251-265; 315-323 <K12>
R:Leffers, H.; Deigaard, K.; Celis, J.E.
Eur. J. Biochem. 230, 447-453, 1995
A:Title: Characterisation of two major cellular poly(UC)-binding human proteins, each
A:Reference number: S65678; MUID:95331278; PMID:7607214
A:Accession: S65678
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-204; 'V', 206-356 <LEFF>
A:Cross-references: EMBL:X78137; NID:g460770; PIDN:CAA55016.1; PID:g460771
A:Experimental source: AMA cells (transformed human amnion cells)
A:Note: submitted to the EMBL Data Library, March 1994
R:Aasheim, H.C.; Loukianova, T.; Beggerdal, A.; Smeland, E.B.
Nucleic Acids Res. 22, 959-964, 1994
A:Title: Tissue specific expression and cDNA structure of a human transcript encoding
A:Reference number: S43489; MUID:94203810; PMID:8152927
A:Accession: S43489
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-298; 'H' <AAS>
A:Cross-references: EMBL:Z29505; NID:g444020; PIDN:CAA82631.1; PID:g444021
C:Genetics:
A:Gene: GDB:HNRPX
A:Cross-references: GDB:344947
C:Keywords: RNA binding

Query Match 7.18; Score 209.5; DB 2; Length 356;

Best Local Similarity 22.8%; Pred. No. 4.5e-06;
Matches 79; Conservative 70; Mismatches 123; Indels 75; Gaps 12;
OY 136 LPRLVPTQFVGAIGKEGATIRNITKOTQSKIDVHRKENAGAEEKSITILSTPGSUA 255
Db 14 LTRILMGKGVSGIIGKGVKRIRESGARINI-----SGNCPERIITLTP--TNA 67
OY 256 AKSILEIMHKEAQDIKT-----EPIKILAHNNFVRLGICGKGRNLAKIEQPT 306
Db 68 IFKAFMIIDLEEDINSMTNSTAASPPVTLRLVVPATCGSLGKGGCKIKIREST 127
OY 307 DTKITISPLQELTYLPERTIVKGVETCAKEEI---MKKIRESYENDIASNLQ- 361
Db 128 GAOVQVA--GMLPNSATERAITAGVPSVTECVKQICLVLETLQSPOGRVMTIPQY 185
OY 362 -AHLIPGLNLNALGFPPTSGMPPT---SGPP----- 390
Db 184 MPASSFVICAGQDCSDAGYPHATHDLESPFLDAYSIOGHIISPLDLAKLNQVAKQ 245
OY 391 -----SMTPTPYQPE-----QSETETVHOF-IPALSVGAIIGKQGHKOL 431
Db 246 SHFAMHGGTGTFAGIDSSPEVKGWASLDASTOTHELTIPNLLIGCIIGROGANINEI 305
OY 432 SRFAGASIKIA-PAPADAKVRNVIITGPPEAQFRAQRIYCKIEE 477
Db 306 KQMSAQIINIANFVSSG--KQVITIGSAASISLAQYLINARLSSE 350

Search completed: April 24, 2003, 16:15:58
Job time : 25 secs

RESULT 15
A44125
high density lipoprotein-binding protein, l10k - human
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1994
C:Revision: 1
R:McKnight, G.L.; Reasoner, J.; Gilbert, T.; Sundquist, K.O.; Hokland, B.; McKernan, P.A.
J. Biol. Chem. 267, 12131-12141, 1992
A:Title: Cloning and expression of a cellular high density lipoprotein-binding protein
A:Reference number: A44125; MUID:92291094; PMID:1318310
A:Accession: A44125
A:Status: Released
A:Molecule type: PEPTIDE
A:Residues: 1-1268 <MCK>
A:Cross-references: GB:M83789; MID:g183891; PIDN:AAA35962.1; PID:g183892
A:Note: Sequence extracted from NCBI backbone (NCBI:106862, NCBI:106863)
C:Superfamily: vigilin

Query Match
Best Local Similarity 23.9%; Pred. No. 4e-05;
Matches 95; Conservative 69; Mismatches 153; Indels 81; Gaps 18;
OY 202 VPTQFVGAIGKEGATIRNITKOTQSKIDVHRKENAGAEEKSITILSTPGTSAACKSIL 261
Db 157 IPKEHHRFVIGKNGEKLODELTKIPIRPDD---PSNQKITGTGEGIEKARHEVL 213
OY 252 EIMHKEAQDIKFTTEIPKILAHNNFVG---RLIGKGRNLAKIEQDTTKITISPLQEL 318
Db 214 LI--SAPQDKAVERLEYEKAPHPFIAGPYNRLV-----EIMQETGRINPP----- 260
OY 319 TLYNPRT-ITVKNVETCAKAEEMKKIKRESYENDIASMNL-----QAHLPGLNLNA 372
Db 261 --PSVNRTEIVFTGEKEQLAQAVARI-KKIYEKKKTKTTIAVEYKKSQHKYVIGPKGNS 317
OY 373 LGLFPPTSG-----MPPPTG-----GPP-----SMTPTPYQFQSETETV-----HO 410
Db 318 LOEILERTGVSEVPSISSEVILRGEKLGQALFEYKANSSTVSVSAAPSILHR 377
OY 411 FIPALSVGAIIGKQGHKOLSRFAGASIKIAPADAKVRNVIITGPPEAQFKAQGR 470
Db 378 F-----IIGKKGONLAKITO-QMPKVHIEFTGED---KITLEGPTEDVNVVQEOI 424
OY 471 YGKIKE-----ENFVSPKREKVELEAHIRVPSFAAGRVIGKGGKTVNELONLSAEVVVPRD 526
Db 425 EGWYKDLINMDIVVEINIDHFKRHLL-----IGSGANINRIIDQYKVSRIIP-- 472

11

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 16:30:00 ; Search time 1.33662 Seconds
(without alignments)
1695.712 Million cell updates/sec

Title: US-09-897-778-176_COPY_41_51
Perfect score: 11
Sequence: 1 FVDCPDESAL 11

Scoring table: OLIGO
Gap 60.0, Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|-----------|---------------------|
| 1 | 11 | 100.0 | 579 | 000425 | 000425 homo sapien |
| 2 | 7 | 63.6 | 100 | 11 Q9D054 | Q9D054 mus musculus |
| 3 | 7 | 63.6 | 576 | 13 Q42254 | Q42254 gallus gall |
| 4 | 7 | 63.6 | 577 | 4 Q9NZ18 | Q9NZ18 homo sapien |
| 5 | 7 | 63.6 | 577 | 11 Q88477 | Q88477 mus musculus |
| 6 | 7 | 63.6 | 579 | 11 Q9CPN8 | Q9CPN8 mus musculus |
| 7 | 7 | 63.6 | 582 | 13 Q9PW80 | Q9PW80 brachydanio |
| 8 | 7 | 63.6 | 593 | 13 Q57526 | Q57526 xenopus lae |
| 9 | 7 | 63.6 | 594 | 13 Q73932 | Q73932 xenopus lae |
| 10 | 6 | 54.5 | 225 | 10 Q8S303 | Q8S303 nicotiana b |
| 11 | 6 | 54.5 | 234 | 16 Q8UGF6 | Q8UGF6 agrobacteri |
| 12 | 6 | 54.5 | 244 | 16 Q98IF4 | Q98IF4 rhizobium l |
| 13 | 6 | 54.5 | 250 | 16 Q8XQ80 | Q8XQ80 ralstonia s |
| 14 | 6 | 54.5 | 310 | 2 P96562 | P96562 amycolatops |
| 15 | 6 | 54.5 | 313 | 17 Q9HP19 | Q9HP19 halobacteri |
| 16 | 6 | 54.5 | 337 | 2 O52816 | O52816 amycolatops |

| | | | | | | |
|----|---|------|------|----|--------|--------------------|
| 17 | 6 | 54.5 | 406 | 2 | O87675 | O87675 amycolatops |
| 18 | 6 | 54.5 | 406 | 2 | O8RN03 | O8RN03 amycolatops |
| 19 | 6 | 54.5 | 460 | 17 | Q9HRL9 | Q9HRL9 halobacteri |
| 20 | 6 | 54.5 | 857 | 13 | P79708 | P79708 chiloscyll |
| 21 | 6 | 54.5 | 2338 | 16 | O8XT19 | O8XT19 ralstonia s |
| 22 | 5 | 45.5 | 88 | 16 | O9ABN7 | O9ABN7 caulobacter |
| 23 | 5 | 45.5 | 101 | 16 | O8YU34 | O8YU34 anabaena sp |
| 24 | 5 | 45.5 | 104 | 2 | O9ZND8 | O9ZND8 thermus the |
| 25 | 5 | 45.5 | 118 | 2 | O85922 | O85922 sphingomona |
| 26 | 5 | 45.5 | 130 | 16 | Q93J47 | Q93J47 streptomyce |
| 27 | 5 | 45.5 | 131 | 10 | O9ATB1 | O9ATB1 brassica ol |
| 28 | 5 | 45.5 | 134 | 10 | O9ATA0 | O9ATA0 brassica ca |
| 29 | 5 | 45.5 | 135 | 10 | O9AXT3 | O9AXT3 brassica na |
| 30 | 5 | 45.5 | 136 | 16 | O9XAE6 | O9XAE6 streptomyce |
| 31 | 5 | 45.5 | 140 | 16 | O9JZU8 | O9JZU8 neisseria m |
| 32 | 5 | 45.5 | 158 | 10 | O9FR89 | O9FR89 brassica na |
| 33 | 5 | 45.5 | 160 | 16 | Q9JW2 | Q9JW2 neisseria m |
| 34 | 5 | 45.5 | 169 | 12 | P87637 | P87637 dhor virus |
| 35 | 5 | 45.5 | 171 | 10 | O944D7 | O944D7 brassica ol |
| 36 | 5 | 45.5 | 182 | 17 | O9YDI3 | O9YDI3 aeropyrum p |
| 37 | 5 | 45.5 | 190 | 16 | O92XI3 | O92XI3 rhizobium m |
| 38 | 5 | 45.5 | 198 | 16 | O66899 | O66899 aquifex aeo |
| 39 | 5 | 45.5 | 207 | 5 | Q9U8F4 | Q9U8F4 schistosoma |
| 40 | 5 | 45.5 | 209 | 16 | Q8XKF0 | Q8XKF0 clostridium |
| 41 | 5 | 45.5 | 211 | 5 | O96303 | O96303 trichomonas |
| 42 | 5 | 45.5 | 211 | 5 | O96305 | O96305 monoceromo |
| 43 | 5 | 45.5 | 215 | 5 | O96304 | O96304 nosema locu |
| 44 | 5 | 45.5 | 215 | 5 | O97437 | O97437 giardia lam |
| 45 | 5 | 45.5 | 217 | 16 | Q9KMY8 | Q9KMY8 vibrio chol |

ALIGNMENTS

RESULT 1

000425 ID 000425 PRELIMINARY; PRT; 579 AA.
AC 000425; TREMBLrel. 04, Created
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative RNA binding protein KOC (KOC).
GN KOC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;
RL Oncogene 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Mueller-Pillasch F., Lacher U., Wallrapp C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; U97188; AAC35208.1; -;
DR EMBL; U97605; AAC09223.1; -;
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH-domain; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0084; KH_TYPE_1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 579 AA; 63720 MW; AE5C3A8EE3C135C5 CRC64;

Query Match 100.0%; Score 11; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 3.le-06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97220007; PubMed=9121465;
RT Ross A.F., Oleynikov Y.S., Kisliouk E.H., Taneja K.L., Singer R.H.;
RL "Characterization of a beta-actin mRNA zipcode-binding protein.";
DR EMBL: AF026527; AAB82295.1;
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type_1.
DR InterPro: IPR00504; RNA_rec_mot.
DR Pfam: PF00013; KH-domain; 4.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00322; KH; 4.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS50084; KH_TYPE_1; 4.
DR PROSITE: PS50102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AAAF2D1D81C8811 CRC64;
Query Match 63.6%; Score 7; DB 13; Length 576;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVDCPDE 7
Db 41 FVDCPDE 47
RESULT 4
Q9NZ18 PRELIMINARY; PRT; 577 AA.
AC Q9NZ18;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE mRNA-binding protein CRDBP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,
RA Panoutsakopoulos G., Kyriazoglou I., Voutzoulis S., Tsiapalis C.M.,
RA Kittas C., Agnantis N., Pandis N.;
RT "Ectopic expression of a KH-domain containing protein, highly
RT homologous to both human IMP-1 and mouse CRD-BP, in benign and
RT malignant mesenchymal tumors.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF198254; AAF37203.1;
DR HSP; P11940; 1CVJ.
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type_1.
DR InterPro: IPR00504; RNA_rec_mot.
DR Pfam: PF00013; KH-domain; 4.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00322; KH; 4.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS50084; KH_TYPE_1; 4.
DR PROSITE: PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;
Query Match 63.6%; Score 7; DB 4; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVDCPDE 7
Db 41 FVDCPDE 47
RESULT 5

QY 1 FVDCPDESVAL 11
Db 41 FVDCPDESVAL 51
RESULT 2
ID Q9D054 PRELIMINARY; PRT; 100 AA.
AC Q9D054;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 261010INIRIK protein.
GN IGF2BP3 OR 261010INIRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=EMBRYO;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL: AK011797; BAB27848.1;
DR MGD: MGI:1890359; IGF2bp3.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
SQ SEQUENCE 100 AA; 11249 MW; 4D871E37EB9D5466 CRC64;
Query Match 63.6%; Score 7; DB 11; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVDCPDE 7
Db 41 FVDCPDE 47
RESULT 3
O42254 PRELIMINARY; PRT; 576 AA.
AC O42254;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Zipcode-binding protein.
GN ZBP1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

088477 ID 088477 PRELIMINARY; PRT; 577 AA.
AC 088477;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE Coding region determinant binding protein (coding region determinant-binding protein).
DE IGF2BP1 OR CRDBP.
GN IGF2BP1 OR CRDBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9211743; PubMed=1559612;
RA Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.:
RT "Control of c-myc mRNA half-life in vitro by a protein capable of binding to a coding region stability determinant.";
RT Genes Dev. 6:642-654(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9415886; PubMed=8114742;
RA Herrick D.J., Ross J.:
RT "The half-life of c-myc mRNA in growing and serum-stimulated cells: influence of the coding and 3' untranslated regions and role of ribosome translocation.";
RT Mol. Cell. Biol. 14:2119-2128(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94179348; PubMed=8132663;
RA Prokipcak R.D., Herrick D.J., Ross J.:
RT "Purification and properties of a protein that binds to the C-terminal coding region of human c-myc mRNA.";
RT J. Biol. Chem. 269:9261-9269(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=9732234; PubMed=9178886;
RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J., Gruppiso P.A., Ross J.:
RT "Developmental regulation of CRD-BP, an RNA-binding protein that stabilizes c-myc mRNA in vitro.";
RT Oncogene 14:1279-1286(1997).
RN [5]
RP SEQUENCE FROM N.A.
RA Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A., Fleisig A.J.:
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G., Brownstein M.J., Bult C., Fletcher C., Carninci P., de Bonaldo M.F., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momabaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.:
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL; AF061569; AAC72743.1; -.

EMBL; AK013940; BAB29071.1; -.
DR HSP; P11940; ICVJ.
DR MGD; MGI:1890357; Igf2bp1.
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type_1.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00013; KH-domain; 4.
DR Pfam: PF00076; Irm; 2.
DR SMART: SM00322; KH; 4.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS00084; KH_TYPE_1; 4.
DR PROSITE: PS0102; RRM; 2.
SQ SEQUENCE 577 AA; 63450 MW; EFBB1AF2FF9F0344 CRC64;
Query Match 63.6%; Score 7; DB 11; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
QY 1 FVDCPDE 7
|||||||
DB 41 FVDCPDE 47
RESULT 6
ID Q9CPN8 PRELIMINARY; PRT; 579 AA.
AC Q9CPN8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 10 days embryo cDNA, RIKEN full-length enriched library.
DE clone:2610036818, full insert sequence (Igf2 mRNA-binding protein 3).
DE IG2BP3 OR 2610101N1RIK OR MIMP3.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G., Brownstein M.J., Bult C., Fletcher C., Carninci P., de Bonaldo M.F., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momabaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.:
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A., Yuasa Y., Takeda M., Okano H.:
RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications for the developing central nervous system.";
RL J. Neurosci. Res. 0:0-0(2001).
DR EMBL; AK011689; BAB27779.1; -.
DR MGD; AB046173; BAB19755.1; -.
DR InterPro: IPR004087; KH_dom.

DR InterPro: IPR004088; KH_type_1.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00013; KH-domain; 4.
 DR Pfam: PF00076; rim; 2.
 DR SMART: SM00322; KH; 4.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50084; KH_type_1; 4.
 DR PROSITE: PS50102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A435B392B7 CRC64;

Query Match 63.6%; Score 7; DB 11; Length 579;
 Best Local Similarity 100.0%; Pred. No. 0.44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVDCPDE 7
 |||||
 Db 41 FVDCPDE 47

RESULT 7

QSPW80 QSPW80 PRELIMINARY; PRT; 582 AA.
 AC QSPW80;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Vg1 RNA binding protein.
 GN DVRLRBP.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang O., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
 RA Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;
 FT "Vg1 RBP intracellular distribution and evolutionarily conserved
 expression suggest multiple roles during development.";
 RL Mech. Dev. 0:0-0(1999).
 DR EMBL: AF161270; AA045610.1; -
 DR ZFIN: ZDB-GENE-000308-1; dvrlrbp.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam: PF00013; KH-domain; 4.
 DR Pfam: PF00076; rim; 2.
 DR SMART: SM00322; KH; 4.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50084; KH_type_1; 4.
 DR PROSITE: PS50102; RRM; 2.
 SQ SEQUENCE 582 AA; 63351 MW; 9DAE63200681B306 CRC64;

Query Match 63.6%; Score 7; DB 13; Length 582;
 Best Local Similarity 100.0%; Pred. No. 0.44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVDCPDE 7
 |||||
 Db 41 FVDCPDE 47

RESULT 8

O57526 O57526 PRELIMINARY; PRT; 593 AA.
 AC O57526;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE KH domain-containing transcription factor B3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92249652; PubMed=1577195;
 RX Pfaff S.L., Taylor W.L.;
 RT "Characterization of a Xenopus oocyte factor that binds to a
 RT developmentally regulated cis-element in the TIIIA gene.";
 RT Dev. Biol. 151:306-316(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Griffin D., Taylor W.L.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
 RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
 RL Genes Dev. 0:0-0(1998).
 DR EMBL: AF042353; AAC18597.1; -
 DR EMBL: AF042353; AAC18597.1; -
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00013; KH-domain; 4.
 DR Pfam: PF00076; rim; 2.
 DR SMART: SM00322; KH; 4.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50084; KH_type_1; 4.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 SQ SEQUENCE 593 AA; 65385 MW; 5A5AB4BA1D55DF7 CRC64;

Query Match 63.6%; Score 7; DB 13; Length 593;
 Best Local Similarity 100.0%; Pred. No. 0.44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVDCPDE 7
 |||||
 Db 41 FVDCPDE 47

RESULT 9

O73932 O73932 PRELIMINARY; PRT; 594 AA.
 AC O73932;
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Vg1 RNA binding protein variant D.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
 RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
 RL Genes Dev. 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98228351; PubMed=9560341;
 RA Deshler J.O., Highett M.I., Abramson T., Schnapp B.J.;
 RT "A highly conserved RNA-binding protein for cytoplasmic mRNA
 RT localization in vertebrates.";
 RL Curr. Biol. 8:489-496(1998).
 DR EMBL: AF064634; AAC18598.1; -
 DR EMBL: AF064634; AAC18598.1; -
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00013; KH-domain; 4.
 DR Pfam: PF00076; rim; 2.

```

DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH_TYPE_1; 4.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 594 AA; 65643 MW; 54CEA7BFF0856DD6 CRC64;

Query Match 63.68; Score 7; DB 13; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.44; 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 1 FVDCPDE 7
Db 41 FVDCPDE 47

RESULT 10
O8S3U3 PRELIMINARY; PRT; 225 AA.
ID Q8S3U3;
AC Q8S3U3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Pf-like protein (Fragment).
OS Nicotiana benthamiana.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21864904; PubMed=11874570;
RA Pearl J.R., Cook G., Feys B.J., Parker J.E., Baulcombe D.C.;
RT "An EDS1 orthologue is required for N-mediated resistance against
RL tobacco mosaic virus."
RL Plant J. 29:569-579(2002).
DR EMBL; AF479624; AAL85346.1; -.
FT NON_TER 1
FT NON_TER 225
SQ SEQUENCE 225 AA; 25674 MW; 1E173D5E15F9860B CRC64;

Query Match 54.5%; Score 6; DB 10; Length 225;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DESWAL 11
Db 134 DESWAL 139

RESULT 11
O8UGF6 PRELIMINARY; PRT; 234 AA.
ID O8UGF6;
AC O8UGF6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Branched-chain amino acid permease.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutavain T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung W., Krespan W., Perry M.,
RA Gordon-Ramm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,

Query Match 54.5%; Score 6; DB 16; Length 244;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DESWAL 11
Db 121 DESWAL 126

RESULT 13
O8XOB0

```

```

RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Fartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AF009072; AAL42095.1; -.
DR EMBL; AF008038; AAK86891.1; -.
KW Complete proteome.
SQ SEQUENCE 234 AA; 24282 MW; 2D6931DA08A13758 CRC64;

```

```

Query Match 54.5%; Score 6; DB 16; Length 234;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DESWAL 11
Db 113 DESWAL 118

```

```

RESULT 12
O98IF4 PRELIMINARY; PRT; 244 AA.
ID Q98IF4;
AC Q98IF4;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein ml2427.
GN ML2427.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF002999; BAB49562.1; -.
DR InterPro; IPR004471; AzlC.
DR Pfam; PF03591; AzlC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 244 AA; 25766 MW; E0DCF859FF3CA77F CRC64;

```

```

Query Match 54.5%; Score 6; DB 16; Length 244;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DESWAL 11
Db 121 DESWAL 126

```

```

RESULT 13
O8XOB0

```

```

ID Q8XQB0 PRELIMINARY; PRT; 250 AA.
AC Q8XQB0;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical transmembrane protein Rsp1376.
GN RSP1376 OR RS02082.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choisine N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646084; CAD18527.1; -
DR ToterPro: IPR004471; Azic.
DR Pfam: PF03591; Azic; 1.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 250 AA; 26490 MW; 685F6FB89089EEC2 CRC64;

Query Match 54.5%; Score 6; DB 16; Length 250;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DESWAL 11
DB 127 DESWAL 132

RESULT 14
P96562 PRELIMINARY; PRT; 310 AA.
ID P96562;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Putative cytochrome P450 16SC2 (Hypothetical hydroxylase D)
DE (Fragment).
GN CYP16SC2.
OS Amycolatopsis orientalis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=31958;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C329.4;
RA Solenberg P.J., Matsushima P., Stack D.R., Wilkie S.C., Thompson R.C.,
RA Baltz R.H.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: U84350; AAB49296.1; -
DR HSSP: Q00441; LOXA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Electron transport; Heme; Hypothetical protein; Membrane;
KW Monooxygenase; Oxidoreductase.
FT NON_TER 1
FT BINDING 260 260 HEME (BY SIMILARITY).
SQ SEQUENCE 310 AA; 34255 MW; 0155572AFA694487 CRC64;

Query Match 54.5%; Score 6; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 5.1;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVDCPD 6
DB 299 FVDCPD 304

---RESULT 15
Q9HP19 PRELIMINARY; PRT; 313 AA.
ID Q9HP19;
AC Q9HP19;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Vng1845C.
GN VNG1845C.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W., Dale H.,
RA Maddocks D.G., Jablonski P.E., Krebs W.P., Angevine C.W., Jung K.-H.,
RA Alam M., Freitas T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Ebnhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE005085; AAG20051.1; -
DR InterPro: IPR000182; GCN5acetyltransf.
DR Pfam: PF00583; Acetyltransf; 1.
KW Complete proteome.
SQ SEQUENCE 313 AA; 34330 MW; 6233C497ED4F8303 CRC64;

Query Match 54.5%; Score 6; DB 17; Length 313;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DESWAL 11
DB 186 DESWAL 191

Search completed: April 16, 2003, 16:36:41
Job time : 3.33662 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2003, 16:11:07 : Search time 38 seconds
(without alignments)
3139.507 Million cell updates/sec

Title: US-09-897-778-176
Perfect score: 2956
Sequence: 1 MNKLYIGNLSENAPSLES.....VKHQOQKALOGPPQSRK 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mmc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|---------------------|
| 1 | 2956 | 100.0 | 579 | 4 000425 | 000425 homo sapien |
| 2 | 2858 | 96.7 | 579 | 11 Q9CPN8 | Q9CPN8 mus musculus |
| 3 | 2483 | 84.0 | 593 | 13 Q57526 | Q57526 xenopus lae |
| 4 | 2461.5 | 83.3 | 594 | 13 Q73932 | Q73932 xenopus lae |
| 5 | 2261.5 | 76.5 | 582 | 13 Q9PW80 | Q9PW80 brachydanio |
| 6 | 2219.5 | 75.1 | 576 | 13 Q42254 | Q42254 gallus gall |
| 7 | 2192 | 74.2 | 577 | 11 Q88477 | Q88477 mus musculus |
| 8 | 2181 | 73.8 | 577 | 4 Q9N218 | Q9N218 homo sapien |
| 9 | 1868 | 63.2 | 556 | 4 Q9Y6M1 | Q9Y6M1 homo sapien |
| 10 | 835.5 | 28.3 | 558 | 5 Q9VZ69 | Q9VZ69 drosophila |
| 11 | 835 | 28.2 | 566 | 5 Q9NGS9 | Q9NGS9 drosophila |
| 12 | 473 | 16.0 | 100 | 11 Q9D054 | Q9D054 mus musculus |
| 13 | 418.5 | 14.2 | 854 | 5 Q21605 | Q21605 caenorhabdi |
| 14 | 262.5 | 8.9 | 774 | 10 Q9LI28 | Q9LI28 oryza sativ |
| 15 | 260.5 | 8.8 | 542 | 10 Q8S7G1 | Q8S7G1 oryza sativ |
| 16 | 242.5 | 8.2 | 557 | 5 Q23487 | Q23487 caenorhabdi |

Q12828 homo sapien
Q91w18 mus musculus
Q9blw0 caenorhabdi.
Q91xf5 arabidopsis
Q9c553 arabidopsis
Q96ae4 homo sapien
Q92946 homo sapien
Q9unw9 homo sapien
Q43267 homo sapien
Q91282 arabidopsis
Q95s29 drosophila
Q91kn6 mus musculus
Q9blw1 caenorhabdi
Q91913 schizosacch
P91632 drosophila
Q96ep6 homo sapien
Q9u982 drosophila
Q9v8h6 drosophila
Q9sr13 arabidopsis
Q17935 caenorhabdi
Q17936 caenorhabdi
Q8uvd9 gallus gall
Q9x171 arabidopsis
Q9sail xenopus lae
Q9asx3 arabidopsis
Q9fmx3 arabidopsis
Q9db01 mus musculus
Q9fmf0 arabidopsis
Q8w5c2 oryza sativ

ALIGNMENTS

RESULT 1

000425 ID 000425 PRELIMINARY; PRT: 579 AA.

AC 000425;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Putative RNA binding protein KOC (KOC).

GN KOC.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;

RL Oncogene 0:0-0(0).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA Mueller-Pillasch F., Lacher U., Wallrapp C.;

RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U97188; AAC35208.1; -.

DR EMBL; U76705; AAC03223.1; -.

DR InterPro; IPR004087; KH_dom.

DR InterPro; IPR004088; KH_type_1.

DR Pfam; PF00013; KH-domain; 4.

DR Pfam; PF00076; rrm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH_type_1; 4.

DR PROSITE; PS50102; RRM; 2.

DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.

SQ SEQUENCE 579 AA; 63720 MW; AE5C3A8EE3C135C5 CRC64;

Query Match 100.0%; Score 2956; DB 4; Length 579;

Best Local Similarity 100.0%; Pred No. 1.7e-187; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKLYIGNLSENAAPS... 60
DB 1 MNKLYIGNLSENAAPS... 60
QY 61 IELHGKPIEVSHSVK... 120
DB 61 IELHGKPIEVSHSVK... 120
QY 121 VNVVYSSKDOARQAL... 180
DB 121 VNVVYSSKDOARQAL... 180
QY 181 ROGSPGVS... 240
DB 181 ROGSPGVS... 240
QY 241 EKSITILSTPGT... 300
DB 241 EKSITILSTPGT... 300
QY 301 KIEQDTRK... 360
DB 301 KIEQDTRK... 360
QY 361 QAHLIPGLN... 420
DB 361 QAHLIPGLN... 420
QY 421 IKGQSHIK... 480
DB 421 IKGQSHIK... 480
QY 481 SPKEVKLEA... 540
DB 481 SPKEVKLEA... 540
QY 541 GHFYACOVA... 579
DB 541 GHFYACOVA... 579

RESULT 2
Q9CPN8
ID Q9CPN8 PRELIMINARY; PRT: 579 AA.
AC Q9CPN8:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE 10 days embryo cDNA, RIKEN full-length enriched library.
DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3).
GN IGFBP3 OR 2610101NIRIK OR MIMP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuenl P., Lewis S., Matsuo Y., Nikaado I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Mattach S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
RA Yuasa Y., Takeda M., Okano H.,
RA "Expression of mouse Igf2 mRNA-binding protein 3 and its implications
RT for the developing central nervous system."
RL J. Neurosci. Res. 0:0-0(2001).
DR EMBL: AK011689; BAB27779.1; -;
DR EMBL: AB046173; BAB19755.1; -;
DR MGD: MGI:1890359; Igfbp3.
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type_1.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00013; KH-domain; 4.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00322; KH; 4.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS00084; KH_TYPE_1; 4.
DR PROSITE: PS0102; RRM; 2.
SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

Query Match 96.7%; Score 2858; DB 11; Length 579;
Best Local Similarity 96.4%; Pred. No. 5.2e-181;
Matches 558; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
QY 1 MNKLYIGNLSENAAPS... 60
DB 1 MNKLYIGNLSENAAPS... 60
QY 61 IELHGKPIEVSHSVK... 120
DB 61 IELHGKPIEVSHSVK... 120
QY 121 VNVVYSSKDOARQAL... 180
DB 121 VNVVYSSKDOARQAL... 180
QY 181 ROGSPGVS... 240
DB 181 ROGSPGVS... 240
QY 241 EKSITILSTPGT... 300
DB 241 EKSITILSTPGT... 300
QY 301 KIEQDTRK... 360
DB 301 KIEQDTRK... 360
QY 361 QAHLIPGLN... 420
DB 361 QAHLIPGLN... 420
QY 421 IKGQSHIK... 480
DB 421 IKGQSHIK... 480
QY 481 SPKEVKLEA... 540
DB 481 SPKEVKLEA... 540
QY 541 GHFYACOVA... 579
DB 541 GHFYACOVA... 579

RESULT 3
ID 057526
AC 16:2

